KINEMATIC ENDPOINT VARIABLES AND PHYSICS BEYOND THE STANDARD MODEL

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Nicholas Scott Eggert

January 2014

KINEMATIC ENDPOINT VARIABLES AND PHYSICS BEYOND THE STANDARD MODEL

Nicholas Scott Eggert, Ph.D.

Cornell University 2014

The Large Hadron Collider (LHC) is currently the world's premier facility for the study of high-energy particle physics. The goal of the Compact Muon Solenoid (CMS) experiment at the LHC is to search for new phenomena in particle physics at the TeV scale. One of the mysteries that data from CMS may solve is the origin of dark matter. It is hoped that dark matter particles will be produced in the proton-proton collisions of the LHC, either by themselves or along with other new particles. This would allow the laws that govern dark matter to be explored.

Collisions containing dark matter particles will pose new challenges for collider physics, necessitating the development of new methods to deal with these problems. One class of methods for addressing these problems is kinematic endpoint variables such as M_{T2} and M_{CT}. A great deal of theoretical work has been done in developing these methods, but they have seldom been applied in experiment. We apply these methods to data taken with CMS to address two problems: mass measurement in underconstrained events and searches for new physics. In doing so, we also provide a new measurement of the top-quark mass and search for new physics in events with two leptons and a momentum imbalance.

BIOGRAPHICAL SKETCH

Nic was born and raised in rural southeastern Wisconsin. He graduated from Horicon High School in Horicon, WI in 2005. He then attended the University of Minnesota in Minneapolis. At the U of M, Nic got involved with particle physics by doing undergraduate research on the NOνA and CMS experiments under Prof. Roger Rusack. Nic's work on CMS included a summer at the CERN laboratory in Geneva, Switzerland. He wrote his undergraduate honors thesis on a measurement of the stopping power of the lead tungstate CMS electromagnetic calorimeter to cosmic-ray muons. While at the University of Minnesota, Nic was awarded several scholarships, including the Alfred O.C. Nier and Edmond G. Franklin scholarships.

Nic graduated Summa Cum Laude from the University of Minnesota with a Bachelor of Science in Physics in the spring of 2009 and immediately started doing research at Cornell University under Prof. Jim Alexander. He started classes and research towards a PhD in physics in August of 2009. After a brief foray working on an x-ray beam size monitor for the CESR Test Accelerator project, Nic returned to working on the CMS experiment under Prof. Alexander for his dissertation research. In 2010, Nic was awarded a Department of Energy Office of Science Graduate Research Fellowship, which funded the remainder of his PhD work. In 2013, he married Christina Cowman, whom he met in a physics class at Minnesota. After graduating, he will remain at Cornell while Christina finishes her PhD in chemistry, working as a Postdoctoral Associate on the Muon $g-2$ experiment.

ACKNOWLEDGMENTS

First of all, I have to thank my advisor, Jim Alexander, for all of the help and guidance he has given me as I worked on my PhD. I'd also like to thank my colleagues at Cornell, especially, Nathan Mirman, Ben Nachman, and Adam Dishaw for their substantial contributions to the top-quark mass measurement. To the Cornell folks at CERN: thank you for being so welcoming in the few weeks a year I spent there.

I'd like to thank the EWKino working group in CMS, especially Sunil Somalwar and Ben Hooberman, for helping me navigate the CMS approval process and for giving suggestions that ultimately made my work better. I also need to thank both the SUSY and Top conveners for facilitating the great work that is done in those groups.

I, of course, need to thank my many other collaborators within CMS. Science on this scale is really a joint effort, and my work would not be possible without all of the work that was put into writing software, building the detector, and the countless critical tasks that I probably don't even know about. I must also thank the LHC team for providing us with collisions and the CERN organization for providing infrastructure and hosting the experiment.

On a more personal level, I want to thank my family and friends who have supported me in this endeavor. First, thanks to my wife, C, for being there with me, in both good times and bad, all the way since freshman year of college. Now that I think about it, I probably wouldn't be here if she hadn't convinced me that I needed to study a little bit. Her work ethic has always been an inspiration to me. I want to thank Alex, Corky, and Matt, whom I've had lunch with almost every day since I started grad school. I've enjoyed our esoteric lunch discussions immensely, and it's been nice to have a group of people with new perspectives to

bounce ideas off. Also, thank you to the many other friends I've made during my time in Ithaca. If I listed everyone individually, people would think I was just trying to pad the length of my dissertation, so I'll just say thank you, everyone. You know who you are.

Finally, I want to thank the US Department of Energy. My research has been supported in part by the Department of Energy Office of Science Graduate Fellowship Program (DOE SCGF). The DOE SCGF Program was made possible in part by the American Recovery and Reinvestment Act of 2009. The DOE SCGF program is administered by the Oak Ridge Institute for Science and Education for the DOE. ORISE is managed by Oak Ridge Associated Universities (ORAU) under DOE contract number DE-AC05-06OR23100.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The world of particle physics is currently best explained by the Standard Model (SM). The SM has been an extraordinary successful theory. It has correctly predicted experimental results thousands of times across 9 orders of magnitude in energy, and has been widely accepted since the 1970s. A great deal of time and effort has been put into experimental searches for deviations from the SM, but scientists have thus far have failed to produce convincing direct evidence that the SM is wrong.

Yet, we suspect that the SM cannot be completely correct. First, it completely fails to incorporate gravity. Despite this, we might suppose that the SM is a very good effective theory at low energies where gravity is not important. However, even this view has its problems. The major problem with the SM is usually referred to as the "hierarchy problem". This problem arises when calculating corrections to the mass of the Higgs boson. We recently learned that the mass of the Higgs boson is around 126 GeV [1, 2]. However, when we calculate the first-order correction to the Higgs boson mass, we find that the correction is equal to the cutoff scale at which new physics appears. If we believe that the SM holds all the way up the the Planck scale, this correction must be on the order of 10^{19} GeV. If the correction is so large, it must mean the the bare Higgs boson mass is comparably large in order to achieve the relatively low Higgs boson mass we observe. Thus, the observed mass is the result of a cancellation over 17 orders of magnitude. While there is no theorem saying that this cannot be the case, it is widely viewed as being "unnatural", and a sign that there must be some as of yet unexplored phenomena below the Planck mass.

Another failure of the SM is its difficulty in explaining dark matter. As-

tronomers have confirmed through a wide variety of methods that most of the mass of the universe is composed of massive particles that do not interact through the strong or electromagnetic forces. The SM does not predict any particle that fits the description of the observed dark matter. Thus we must hypothesize that there is more to particle physics than the Standard Model.

One of the leading candidate theories for describing physics beyond the SM is called supersymmetry, commonly referred to as SUSY. This theory introduces extra particles which cancel these huge contributions to the Higgs boson mass, leaving a more natural theory. There is also a large region of SUSY parameter space in which lightest supersymmetric particle (LSP) is stable, neutral, and interacts only through the weak force. This predicted particle is a perfect match for the expected properties of dark matter. The theory behind SUSY is discussed in Chapter 2.

In order to study the regime in which these problems occur, we need to study the interactions of particles at high energies. The Large Hadron Collider (LHC) is currently the world's premier facility for these studies. The LHC collides proton bunches together at center of mass energies of around 10 TeV. The results of these collisions are recorded by four detectors: CMS, ATLAS, LHCb, and ALICE. The LHC and the CMS detector are described in more detail in Chapter 3. The detectors reconstruct all of the particles produced by the collisions, making it possible to understand the physics that governs the collisions. The techniques and algorithms used for reconstructing particles are presented in Chapter 4.

One of the major difficulties in the search for new physics, SUSY or otherwise, arises from the fact that any new theory must account for dark matter. By the very nature of dark matter, it is not easily detected. Its presence can often be inferred, but it is not directly visible to a detector like CMS. To add to the difficulty, most theories predict that dark matter particles will appear in pairs. This leads to events

that are kinematically underconstrained; there is not enough information to fully reconstruct the event.

To overcome this challenge, a new class of methods has been developed. These methods use kinematic variables that exhibit upper bounds or endpoint structures. For these variables, the location of the endpoint contains information about the masses of the particles in the events. A great deal of theoretical work on these variables has been done in the last 10 years. Chapter 5 provides a brief review of important results. The experimental community has been somewhat slow to take advantage of this work. In particular, it has not been clear how well kinematic endpoint methods would work on real data taken with a non-ideal detector. Chapter 6 remedies this by using kinematic endpoint variables to measure particle masses in kinematically underconstrained SM events.

Chapter 7 shows another way in which kinematic endpoint variables can be used: as discriminating variables in searches for new physics. We search for underconstrained new physics processes that differ kinematically from SM processes only by the masses of particles. Since kinematic endpoint variables are directly sensitive to the masses of particles, they provide a good discriminator against the SM backgrounds. The analysis focusses on weakly-produced particles, as the lack of hadronic activity gives very well-measured kinematic variables.

When and if physics beyond the Standard Model is discovered, kinematic endpoint variables will be an important technique for studying new processes and hopefully resolving both the hierarchy and dark matter problems. We show that these techniques can be usefully applied to real data. In the process, we also provide a novel measurement of the top-quark mass and search for weakly-produced new physics.

CHAPTER 2

THEORY

Let us begin with a brief introduction to the Standard Model (SM). We will then explore in some detail one of the most promising theories for physics beyond the SM: supersymmetry (SUSY). We also discuss the experimental consequences of SUSY.

2.1 The Standard Model

The Standard Model is the currently accepted theory describing particle physics. This theory describes subatomic physics in terms of quantum field theory. The particles that make up matter are described as spin 1/2 quantum fields. The behavior of these fields is specified by a Lagrangian. The interactions of these particles come about by requiring that the Lagrangian be invariant under various gauge transformations. By simply defining the $SU(3) \times SU(2)_L \times U(1)$ gauge structure, as well as the representations of the matter particles under these gauge groups, we almost completely specify the fundamental physics describing subatomic particles.

The "almost" has been the subject of most of the work in particle physics for the last 40 years. In order for the Standard Model to be self-consistent, a scalar field with a non-zero expectation value must be added to the theory. This prevents unitarity violation and provides a mechanism for generating the observed masses of the W and Z bosons, which the theory would otherwise predict to be zero. The recent observation of the Higgs boson [1, 2] at the Large Hadron Collider confirms the presence of this scalar field, or at least something like it.

Despite its success, the Standard Model suffers from at least two flaws. The first is in its prediction of the mass of the Higgs boson. When calculating the first-order loop correction to the Higgs boson mass in the Standard Model, one finds that the

correction is on the order of some large cut-off scale. This cut-off scale is either the Planck scale (10^{19} GeV) or the scale at which physics beyond the Standard Model becomes important, if one postulates that such physics exists. This calculation would indicate that the first-order correction is orders of magnitude larger than the Higgs boson mass itself. While this is technically allowed in the theory, it is not considered to be natural, and is taken as evidence of physics beyond the Standard Model. This is referred to as the hierarchy problem.

The second short-coming of the Standard Model is that it does not predict any particle which could account for the dark matter observed by astronomers. A number of considerations make a weakly-interacting massive particle a prime candidate for dark matter, but no such particle is predicted by the Standard Model. Again, this points towards the presence of physics beyond the Standard Model.

2.2 A Basic Supersymmetric Model

One of the leading candidates for describing physics beyond the Standard Model is supersymmetry (SUSY). SUSY provides both a solution to the hierarchy problem, and, in some models, a dark matter candidate. In the following sections, we will give an introduction to the supersymmetric Standard Model and discuss some of its properties and implications.

The basic idea of supersymmetry is to impose a symmetry on the Lagrangian which turns fermions into bosons and vice versa. This idea was not originally introduced to solve the hierarchy problem, but it turns out to do a remarkably good job, canceling the divergent contributions to the Higgs boson mass to all orders in perturbation theory.

We begin by creating a simple and generic supersymmetric model, then move to a more realistic model which obeys the symmetries of the Standard Model.

2.2.1 Chiral Supermultiplets

We first consider the simplest possible Lagrangian upon which we can impose supersymmetry: a single complex scalar field ϕ and a two-component Weyl fermion ψ, both massless. This is referred to as the Wess-Zumino model [3].

$$\mathcal{L} = -\partial^\mu \phi^* \partial_\mu \phi + i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi \tag{2.1}$$

Supersymmetry is a continuous symmetry. The supersymmetry transformation transforms scalar and fermion fields into linear combinations of scalar and fermion fields. We call a pair of fields that transform into each other under a supersymmetry transformation a *supermultiplet*. Since the supersymmetry transformation is continuous, we can consider an infinitesimal transformation:

$$\phi \to \phi + \delta\phi \tag{2.2a}$$

$$\psi \to \psi + \delta\psi. \tag{2.2b}$$

We now want to find $\delta\phi \propto \psi$ and $\delta\psi \propto \phi$ such that the Wess-Zumino Lagrangian is invariant under the transformations in Equation 2.2. A little manipulation shows that the transformations which leave the Lagrangian unchanged are

$$\delta\phi = \epsilon\psi \tag{2.3a}$$

$$\delta\phi^* = \epsilon^\dagger \psi^\dagger \tag{2.3b}$$

$$\delta\psi = -i\sigma^\mu \epsilon^\dagger \partial_\mu \phi \tag{2.3c}$$

$$\delta\psi^\dagger = i\epsilon\sigma^\mu \partial_\mu \phi^*, \tag{2.3d}$$

where ϵ is an infinitesimal two-component Weyl spinor which parameterizes the transformation.

It turns out that in order to have a consistent theory, we must make a slight modification. We want the supersymmetry transformations to obey the same algebra regardless of whether they are applied on the scalar or fermionic fields. One of the requirements for this is that the operator formed by the commutator of two different supersymmetry transformations must behave in the same way when applied to either field. With the theory just described, this is not the case. Applying the commutator of two transformations to each of the fields gives

$$[\delta_{\epsilon_1}, \delta_{\epsilon_2}]\phi = (\delta_{\epsilon_1}\delta_{\epsilon_2} - \delta_{\epsilon_2}\delta_{\epsilon_1})\phi = i(-\epsilon_1\sigma^\mu\epsilon_2^\dagger + \epsilon_2\sigma^\mu\epsilon_1^\dagger)\partial_\mu\phi \tag{2.4a}$$

$$[\delta_{\epsilon_1}, \delta_{\epsilon_2}]\psi = (\delta_{\epsilon_1}\delta_{\epsilon_2} - \delta_{\epsilon_2}\delta_{\epsilon_1})\psi = i(-\epsilon_1\sigma^\mu\epsilon_2^\dagger + \epsilon_2\sigma^\mu\epsilon_1^\dagger)\partial_\mu\psi + i\epsilon_1\epsilon_2^\dagger\bar{\sigma}^\mu\partial_\mu\phi - i\epsilon_2\epsilon_1^\dagger\bar{\sigma}^\mu\partial_\mu\psi. \tag{2.4b}$$

We would like to have Equation 2.4a equal to Equation 2.4b. To attain this, we can add a massless, non-propagating, complex scalar field to the Lagrangian. We call this the *auxiliary* field. With this extra field F, the Lagrangian becomes

$$\mathcal{L} = -\partial^\mu\phi^*\partial_\mu\phi + i\psi^\dagger\bar{\sigma}^\mu\partial_\mu\psi + F^*F. \tag{2.5}$$

Note that the Euler-Lagrange equation for the auxiliary field gives $F = F^* = 0$.

To keep invariance under supersymmetry transformations and keep the algebra consistent, we define the supersymmetry transformation for F and modify slightly the transformation for ψ. The new transformations are

$$\delta\phi = \epsilon\psi \tag{2.6a}$$

$$\delta\phi^* = \epsilon^\dagger\psi^\dagger \tag{2.6b}$$

$$\delta\psi = -i\sigma^\mu\epsilon^\dagger\partial_\mu\phi + \epsilon F \tag{2.6c}$$

$$\delta\psi^\dagger = i\epsilon\sigma^\mu\partial_\mu\phi^* + \epsilon^\dagger F^* \tag{2.6d}$$

$$\delta F = i\epsilon^\dagger\bar{\sigma}^\mu\partial_\mu\psi \tag{2.6e}$$

$$\delta F^* = i\partial_\mu\psi^\dagger\bar{\sigma}^\mu\epsilon. \tag{2.6f}$$

With these transformations, Equation 2.4a applies for all three fields and we have a consistent theory.

We now turn to interactions in the Wess-Zumino model. First we introduce the possibility of more than one supermultiplet. Each supermultiplet has the same Lagrangian as Equation 2.5. Our total Lagrangian is just a sum over the different supermultiplets i.

$$\mathcal{L}_{\mathrm{WZ}} = -\partial^\mu \phi^{i*} \partial_\mu \phi_i + i\psi^{i\dagger} \bar{\sigma}^\mu \partial_\mu \psi_i + F^{i*} F_i. \tag{2.7}$$

Here and throughout the chapter, we use the Einstein notation, in which repeated indices are implicitly summed over.

Next, we add interaction terms. The terms allowed to be present in the interaction Lagrangian are severely constrained by the requirement that they respect supersymmetry. Only two interaction terms are allowed:

$$\mathcal{L}_{\mathrm{int}} = -\frac{1}{2} W^{ij} \psi_i \psi_j + W^i F_i, \tag{2.8}$$

where W^{ij} and W^i are polynomials of degree 1 and 2 of the scalar field ϕ_i. Note that they cannot depend on the adjunct of the scalar fields, ϕ^{i*}, as this would introduce terms that cannot be made invariant under supersymmetry.

While it is not *a priori* apparent, imposing supersymmetry upon the interaction Lagrangian forces particular forms on W^{ij} and W^i. Because of the forms they must take, W^{ij} and W^i can be related to a quantity called the superpotential W.

$$W = \frac{1}{2} M^{ij} \phi_i \phi_j + \frac{1}{6} y^{ijk} \phi_i \phi_j \phi_k \tag{2.9a}$$

$$W^i = \frac{\delta W}{\delta \phi_i} = M^{ij} \phi_j + \frac{1}{2} y^{ijk} \phi_j \phi_k \tag{2.9b}$$

$$W^{ij} = \frac{\delta^2 W}{\delta \phi_i \delta \phi_j} = M^{ij} + y^{ijk} \phi_k \tag{2.9c}$$

We can now write down our total Lagrangian.

$$\mathcal{L} = -\partial^\mu \phi^{i*} \partial_\mu \phi_i + i\psi^{i\dagger} \bar{\sigma}^\mu \partial_\mu \psi_i + F^{i*} F_i - \frac{1}{2} W^{ij} \psi_i \psi_j + W^i F_i - \frac{1}{2} W^*_{ij} \psi^{i\dagger} \psi^{j\dagger} + W^{i*} F^*_i \tag{2.10}$$

We can use the Euler-Lagrange equations to find that $F_i = -W^*_i$, and $F^{i*} = -W^i$. Filling in these expressions for W, we get

$$\begin{aligned}
\mathcal{L} =& -\partial^\mu \phi^{i*} \partial_\mu \phi_i + i\psi^{i\dagger} \bar{\sigma}^\mu \partial_\mu \psi_i + \frac{1}{2} W^{ij} \psi_i \psi_j + \frac{1}{2} W^*_{ij} \psi^{i\dagger} \psi^{j\dagger} - W^{i*} W_i & (2.11)\\
=& -\partial^\mu \phi^{i*} \partial_\mu \phi_i + i\psi^{i\dagger} \bar{\sigma}^\mu \partial_\mu \psi_i - \frac{1}{2} \left(M^{ij} \psi_i \psi_j + y^{ijk} \psi_i \psi_j \phi_k + \text{c.c.} \right) & (2.12)\\
& - M^*_{ik} M^{kj} \phi^{i*} \phi_j - \frac{1}{2} \left(M^{in} y^*_{jkn} \phi_i \phi^{j*} \phi^{k*} + \text{c.c} \right) \\
& - \frac{1}{4} y^{ijn} y^*_{kln} \phi_i \phi_j \phi^{k*} \phi^{l*}.
\end{aligned}$$

Note that the last term in Equation 2.11 and the corresponding last two lines in Equation 2.12 are usually referred to as F-terms, since they come from the auxiliary field F.

We see that our Lagrangian now includes mass terms for the scalar and fermion fields, as well as an interaction between the fermions and scalars and self-interactions of the scalars. If we were to derive the linearized equations of motion for the fields and do a bit of manipulation, we would find that the scalars and fermions have the same mass matrix. Since this is not the case in nature, we will eventually need to introduce supersymmetry-breaking terms.

2.2.2 Gauge Supermultiplets

We now turn to gauge supermultiplets for a generic gauge group. These are composed of the spin 1 gauge fields A^a_μ and a corresponding spin 1/2 gaugino λ^a for each gauge field. This gaugino field transforms under an infinitesimal gauge transformation parameterized by Λ^a as

$$\lambda^a \to \lambda^a + g f^{abc} \lambda^b \Lambda^c, \tag{2.13}$$

where g is the gauge coupling constant and f^{abc} are the gauge group structure constants.

As was the case for the chiral supermultiplets, it turns out that we need to introduce a fermionic auxiliary field D^a to keep the theory consistent. This D^a field transforms under gauge transformations in the same way as λ^a.

The Lagrangian for the gauge supermultiplet is

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu a} + i\lambda^{a\dagger}\bar{\sigma}^\mu\nabla_\mu\lambda^a + \frac{1}{2}D^aD^a, \tag{2.14}$$

where

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu \tag{2.15}$$

and

$$\nabla_\mu\lambda^a = \partial_\mu\lambda^a + gf^{abc}A^b_\mu\lambda^c \tag{2.16}$$

is the gauge covariant derivative.

The field transformations that make the gauge Lagrangian invariant under supersymmetry are

$$\delta A^a_\mu = -\frac{1}{\sqrt{2}}\left(\epsilon^\dagger\bar{\sigma}_\mu\lambda^a + \lambda^{a\dagger}\bar{\sigma}_\mu\epsilon\right) \tag{2.17a}$$

$$\delta\lambda^a = \frac{i}{2\sqrt{2}}\sigma^\mu\bar{\sigma}^\nu\epsilon F^a_{\mu\nu} + \frac{1}{\sqrt{2}}\epsilon D^a \tag{2.17b}$$

$$\delta D^a = \frac{i}{\sqrt{2}}\left(-\epsilon^\dagger\bar{\sigma}^\mu\nabla_\mu\lambda^a + \nabla_\mu\lambda^{a\dagger}\bar{\sigma}^\mu\epsilon\right) \tag{2.17c}$$

2.2.3 The Interacting Wess-Zumino Model

We now want to combine our gauge supermultiplet with a chiral supermultiplet to produce a gauge invariant theory of matter particles. Most of the work here is done by the usual procedure of swapping derivatives in the chiral supermultiplet

Lagrangian for their gauge covariant versions:

$$\begin{aligned}\nabla_\mu \phi_i &= \partial_\mu \phi_i - igA^a_\mu (T^a \phi)_i \\ \nabla_\mu \phi^{i*} &= \partial_\mu \phi^{i*} + igA^a_\mu (\phi^* T^a)^i \\ \nabla_\mu \psi_i &= \partial_\mu \psi_i - igA^a_\mu (T^a \psi)_i,\end{aligned} \tag{2.18}$$

where T^a are the generators of the gauge group.

There are also a few allowed couplings between the chiral fields and the gauginos and auxilliary fields that must be added:

$$(\phi^* T^a \psi)\, \lambda^a \tag{2.19a}$$

$$\lambda^{a\dagger} \left(\psi^\dagger T^a \phi\right) \tag{2.19b}$$

$$(\phi^* T^a \phi)\, D^a. \tag{2.19c}$$

To make the whole mess invariant under sypersymmetry, we need to modify slightly the supersymmetry transformations of the chiral fields:

$$\delta \phi_i = \epsilon \psi_i \tag{2.20a}$$

$$\delta \psi_i = -i\sigma^\mu \epsilon^\dagger \nabla_\mu \phi_i + \epsilon F_i \tag{2.20b}$$

$$\delta F_i = i\epsilon^\dagger \bar{\sigma}^\mu \nabla_\mu \psi_i + \sqrt{2} g (T^a \phi)_i \epsilon^\dagger \lambda^{a\dagger}. \tag{2.20c}$$

Similar transformations apply for the charge conjugates. With a little bit of work, the supersymmetry requirement fixes the coefficients for the couplings in Equation 2.19.

$$\mathcal{L} = \mathcal{L}_{\text{chiral}} + \mathcal{L}_{\text{gauge}} - \sqrt{2} g \left(\phi^* T^a \psi\right) \lambda^a - \sqrt{2} g \lambda^{a\dagger} \left(\psi^\dagger T^a \phi\right) + g \left(\phi^* T^a \phi\right) D^a \tag{2.21}$$

Finally, we insert the solution for D^a from the Euler-Lagrange equations, $D^a =$

$-g(\phi^* T^a \phi)$, giving us the whole Lagrangian:

$$\begin{aligned}\mathcal{L} = &- \nabla^\mu \phi^{i*} \nabla_\mu \phi_i + i\psi^{i\dagger} \bar{\sigma}^\mu \nabla_\mu \psi_i - \frac{1}{2}\left(M^{ij}\psi_i\psi_j + y^{ijk}\psi_i\psi_j\phi_k + \text{c.c.}\right) \\ &- M^*_{ik} M^{kj} \phi^{i*}\phi_j - \frac{1}{2}\left(M^{in} y^*_{jkn} \phi_i \phi^{j*} \phi^{k*} + \text{c.c}\right) \\ &- \frac{1}{4} y^{ijn} y^*_{kln} \phi_i \phi_j \phi^{k*} \phi^{l*} \\ &- \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} + i\lambda^{a\dagger} \bar{\sigma}^\mu \nabla_\mu \lambda^a \\ &- \sqrt{2} g \left(\phi^* T^a \psi\right) \lambda^a - \sqrt{2} g \lambda^{a\dagger} \left(\psi^\dagger T^a \phi\right) + \frac{1}{2} g^2 \left(\phi^* T^a \phi\right)^2,\end{aligned} \tag{2.22}$$

where c.c. stands for charge conjugate. We must of course, sum over the gauge fields A^a_μ and their corresponding generators T^a. Note that the last term in this expression is called the D-term, for reasons similar to those that named the F-term.

2.2.4 Supersymmetry-breaking

As mentioned earlier, the Wess-Zumino theory has identical sfermion and fermion mass matrices. If we are to create a theory that might describe the real world, we must break the supersymmetry. Currently, there is no completely satisfactory mechanism for this, so we introduce effective SUSY-breaking terms to the Lagrangian. We are somewhat constrained in what terms we can add by the requirement that our theory still cancels the quadric divergences in the Higgs boson mass. This restricts us to terms with coupling constants having positive mass dimension, called soft SUSY-breaking terms. The allowed terms, minus a few that are not used in the Supersymmetric Standard Model, are

$$\mathcal{L}_{\text{soft}} = -\left(\frac{1}{2} M_a \lambda^a \lambda^a + \frac{1}{6} a^{ijk} \phi_i \phi_j \phi_k + \frac{1}{2} b^{ij} \phi_i \phi_j + \text{c.c.}\right) - \left(m^2\right)^i{}_j \phi^{j*} \phi_i \tag{2.23}$$

Inspecting the terms in Equation 2.23, we see that we have gained mass terms for the gauginos. We also modified the trilinear sfermion couplings and the sfermion masses.

2.3 The Minimal Supersymmetric Standard Model

To construct the Minimal Supersymmetric Standard Model (MSSM), we start by assigning Standard Model matter particles to chiral supermultiplets. We then impose the $\mathrm{SU(3)_C \times SU(2)_L \times U(1)_Y}$ gauge symmetry of the SM onto our theory, generating gauge supermultiplets. Then we add a Higgs sector to generate masses for the electroweak gauge bosons and the fermions. Finally, we add soft SUSY breaking.

2.3.1 Particle Content

We assign each of the SM fermions to its own chiral supermultiplet, which means that each fermion gains a scalar partner. We follow the notation of Ref. [4], using Q_i to denote the left handed doublet of quarks of the ith family. That is:

$$Q_1 = \begin{pmatrix} \mathrm{u_L} \\ \mathrm{d_L} \end{pmatrix}, \qquad Q_2 = \begin{pmatrix} \mathrm{c_L} \\ \mathrm{s_L} \end{pmatrix}, \qquad Q_3 = \begin{pmatrix} \mathrm{t_L} \\ \mathrm{b_L} \end{pmatrix}. \tag{2.24}$$

We use $\bar{u}_i$ and $\bar{d}_i$ to denote the right-handed singlet quarks of the up and down type, respectively.

Similarly, we use L_i to denote the left-handed doublets of leptons and $\bar{e}_i$ the right-handed singlets. Each of these, of course, has its scalar partner, denoted with a tilde, i.e. $\tilde{Q}_i$, $\tilde{\bar{d}}_i$.

We also at this point introduce $\mathrm{SU(2)_L}$ Higgs doublets, which we will need later for electroweak symmetry breaking. It is necessary to have two because, as mentioned earlier, the superpotential is not allowed to be a function of the adjoints of scalar fields. Thus, where we would use something like $H^\dagger$ in the Standard Model to give mass to the down-type quarks, in the MSSM, we must introduce a second Higgs doublet with weak hypercharge $Y = -1$. The two Higgs doublets are

$$H_u = \begin{pmatrix} \mathrm{H}_u^+ \\ \mathrm{H}_u^0 \end{pmatrix}, \qquad H_d = \begin{pmatrix} \mathrm{H}_d^0 \\ \mathrm{H}_d^- \end{pmatrix}. \tag{2.25}$$

The Higgs field are the scalar parts of supermultiplets, so they must also have fermionic partners. We call these the Higgsinos, denoted $\tilde{H}_u$ and $\tilde{H}_d$.

Having constructed the chiral supermultiplets, it is fairly straightforward to add the gauge symmetries of the Standard Model. We change the derivatives in the Wess-Zumino Lagrangian to be covariant in all three gauge groups. Also need to add the kinetic and gaugino-(s)fermion interactions for each gauge field. As we might expect, we end up with 8 SU(3) gauge fields, 3 SU(2) gauge fields, and 1 U(1) gauge field. We also get their spin 1/2 partners, the 8 gluinos, 3 winos, and 1 bino.

2.3.2 Lagrangian

In the end, we have a Lagrangian that looks like Equation 2.22. We sum i and j over all of the fermions. The covariant derivatives are covariant under all three gauge groups. The gauge boson and gaugino kinetic terms, along with the chiral-gaugino interactions, are also summed over all three gauge fields.

All that is left to do is to specify the masses and Yukawa couplings by specifying the superpotential. We take M^{ij} to be zero for most of the chiral fields, as we want to generate the masses dynamically through electroweak symmetry breaking. The exceptions are the Higgs supermultiplets, which we allow to have inherent mass. The superpotential we end up with is

$$W = \bar{u}\mathbf{y}_u Q H_u - \bar{d}\mathbf{y}_d Q H_d - \bar{e}\mathbf{y}_e L H_d + \mu H_u H_d, \tag{2.26}$$

where $\mathbf{y}_x$ are 3×3 Yukawa matrices. Since the superpotential is only a function of the scalar fields, we leave off the tildes to reduce clutter.

We must now add soft SUSY-breaking to the MSSM. We do this by adding all

allowed terms of the form in Equation 2.23.

$$\begin{aligned}\mathcal{L}_{\text{soft}} = &-\frac{1}{2}\left(M_3\widetilde{\mathrm{g}}\widetilde{\mathrm{g}} + M_2\widetilde{\mathrm{W}}\widetilde{\mathrm{W}} + M_1\widetilde{\mathrm{B}}\widetilde{\mathrm{B}} + \text{c.c.}\right) \\ &- \left(\tilde{\bar{u}}\mathbf{a}_u\tilde{Q}H_u - \tilde{\bar{d}}\mathbf{a}_d\tilde{Q}H_d - \tilde{\bar{e}}\mathbf{a}_e\tilde{L}H_d + \text{c.c.}\right) \\ &- \tilde{Q}^\dagger\mathbf{m}_q^2\tilde{Q} - \tilde{L}^\dagger\mathbf{m}_L^2\tilde{L} - \tilde{\bar{u}}\mathbf{m}_{\bar{u}}^2\tilde{\bar{u}}^\dagger - \tilde{\bar{d}}\mathbf{m}_{\bar{d}}^2\tilde{\bar{d}}^\dagger - \tilde{\bar{e}}\mathbf{m}_{\bar{e}}^2\tilde{\bar{e}}^\dagger \\ &- m_{H_u}^2H_u^*H_u - m_{H_d}^2H_d^*H_d - (bH_uH_d + \text{c.c.})\end{aligned} \tag{2.27}$$

This introduces mass terms for the gauginos as well as all of the sfermions and Higgs bosons. It also adds couplings between the right and left-handed sfermions through the Higgs boson. In principle the soft SUSY-breaking term introduces 105 new parameters to the MSSM [5]. However, many of these parameters introduce flavor mixing or CP violation, and thus they are considerably constrained by experiment. Examples of these experimental constraints include searches for $\mu \to \mathrm{e}\gamma$ and measurements of $\mathrm{D}^0\overline{\mathrm{D}}^0$ and $\mathrm{K}^0\overline{\mathrm{K}}^0$ mixing.

One way of working within these constraints is to assume universal SUSY breaking. This assumes that the $\mathbf{m}^2$ matrices are diagonal and that the $\mathbf{a}$ matrices are proportional to the Yukawa matrices. We also assume that the new parameters introduce no new complex phases. Under these assumptions, b and μ can be taken to be real through some allowed phase rotations. We then assume that M_1, M_2, M_3 and the constants of proportionality of the $\mathbf{a}$ matrices to the Yukawa matrices are real.

2.4 Particle Masses in the MSSM

In the MSSM, mass terms come from a wide variety of sources including the superpotential, electroweak symmetry breaking, and SUSY-breaking terms. We start with electroweak symmetry breaking, which gives mass to the Standard Model

particles. We then discuss the masses of the electroweak gauginos and first and second-generation sleptons, as these are relevant to the analysis in Chapter 7.

2.4.1 Electroweak Symmetry Breaking

From the Lagrangian we have built up so far, the scalar potential of the Higgs fields is

$$\begin{aligned} V = &|\mu|^2 \left(|H_u^0|^2 + |H_u^+|^2 + |H_d^0|^2 + |H_d^-|^2\right) \\ &+ m_{H_u}^2 \left(|H_u^0|^2 + |H_u^+|^2\right) + m_{H_d}^2 \left(|H_d^0|^2 + |H_d^-|^2\right) \\ &+ b\left(H_u^+ H_d^- - H_u^0 H_d^0 + \text{c.c.}\right) \\ &+ \frac{1}{2}\left(g^2 + g'^2\right)\left(|H_u^0|^2 + |H_u^+|^2 - |H_d^0|^2 - |H_d^-|^2\right)^2 \\ &+ \frac{1}{2} g^2 |H_u^+ H_d^{0*} + H_u^0 H_d^{-*}|^2. \end{aligned} \tag{2.28}$$

The terms in the first line come from F-terms. Lines 2 and 3 come from $\mathcal{L}_{\text{soft}}$, and the final two lines come from the D-terms, slightly rearranged.

Since the Lagrangian is invariant under gauge transformations, we are free to make such a transformation so that $H_u^+ = 0$. At a minimum in the potential, where $\partial V/\partial H_u^+ = 0$ we can see also that $H_d^- = 0$. After making this transformation and grouping some terms, our scalar potential is

$$\begin{aligned} V = &\left(|\mu|^2 + m_{H_u}^2\right)|H_u^0|^2 + \left(|\mu|^2 + m_{H_d}^2\right)|H_d^0|^2 \\ &- b H_u^0 H_d^0 + \frac{1}{8}\left(g^2 + g'^2\right)\left(|H_u^0|^2 - |H_d^0|^2\right)^2. \end{aligned} \tag{2.29}$$

Since the term containing b is the only term that depends on the phase of the fields, we can choose the phase of the field to absorb any phase in b. Therefore, we take b to be real and ≥ 0.

Provided certain conditions are met, this potential provides a stable minimum away from 0. We set the expectation values of the fields at this minimum to

$\langle H_u^0 \rangle = v_u$, $\langle H_d^0 \rangle = v_d$, the vacuum expectation values (VEV). Since both of the VEVs contribute mass terms to the vector bosons, they are related to the SM VEV by

$$v_u^2 + v_d^2 = v_{\mathrm{SM}}^2 = \frac{2m_{\mathrm{Z}}^2}{g^2 + g'^2}. \tag{2.30}$$

We define the ratio of the VEVs to be

$$\tan\beta = \frac{v_u}{v_d}. \tag{2.31}$$

At the minimum of the potential, we have the conditions

$$\frac{\partial V}{\partial H_u^0} = \frac{\partial V}{\partial H_d^0} = 0, \tag{2.32}$$

so we can obtain relations between some of the parameters in the potential,

$$m_{H_u}^2 + |\mu|^2 - b\cot\beta - \frac{m_{\mathrm{Z}}^2}{2}\cos 2\beta = 0 \tag{2.33a}$$

$$m_{H_d}^2 + |\mu|^2 - b\tan\beta + \frac{m_{\mathrm{Z}}^2}{2}\cos 2\beta = 0, \tag{2.33b}$$

Given these 5 unknowns and 2 equations, we can completely specify the Higgs sector by 3 parameters. These are usually chosen to be β, $m_{H_u}^2$, and $m_{H_d}^2$. We also do not know the sign of μ, since only its magnitude enters these calculations.

2.4.2 Electroweak Gaugino Masses

Since the electroweak gauginos and higgsinos share the same quantum numbers, they will, in general, mix. First, consider the neutral fields. We can write down the vector of these particles in the gauge eigenbasis as $\psi^0 = (\widetilde{\mathrm{B}}, \widetilde{\mathrm{W}}, \tilde{\mathrm{H}}_d^0, \tilde{\mathrm{H}}_u^0)$. Collecting all of the mass terms for these fields, we find the mass term in the gauge eigenbasis:

$$\mathcal{L} = -\frac{1}{2}\left(\psi^0\right)^{\mathrm{T}} \mathbf{M}_N \psi^0, \tag{2.34}$$

where

$$\mathbf{M}_N = \begin{pmatrix} M_1 & 0 & -g'v_d/\sqrt{2} & g'v_u/\sqrt{2} \\ 0 & M_2 & gv_d/\sqrt{2} & -gv_u/\sqrt{2} \\ -g'v_d/\sqrt{2} & gv_d/\sqrt{2} & 0 & -\mu \\ g'v_u/\sqrt{2} & -gv_u/\sqrt{2} & -\mu & 0 \end{pmatrix}. \tag{2.35}$$

The terms in the upper-left quadrant come directly from $\mathcal{L}_{\text{soft}}$, and those in the lower-right come from the superpotential. The off-diagonal quadrants come from electroweak symmetry breaking through couplings of the form $-\sqrt{2}g(\phi^* T^a \psi)\lambda^a$ as in Equation 2.21, where ϕ is replaced by the Higgs VEV, ψ by a Higgsino and λ by $\widetilde{\mathrm{B}}$ or $\widetilde{\mathrm{W}}$. We can re-write this matrix in a sightly more illuminating form:

$$\mathbf{M}_N = \begin{pmatrix} M_1 & 0 & -\mathrm{c}_\beta \mathrm{s}_\mathrm{W} m_\mathrm{Z} & \mathrm{s}_\beta \mathrm{s}_\mathrm{W} m_\mathrm{Z} \\ 0 & M_2 & \mathrm{c}_\beta \mathrm{c}_\mathrm{W} m_\mathrm{Z} & -\mathrm{s}_\beta \mathrm{c}_\mathrm{W} m_\mathrm{Z} \\ -\mathrm{c}_\beta \mathrm{s}_\mathrm{W} m_\mathrm{Z} & \mathrm{c}_\beta \mathrm{c}_\mathrm{W} m_\mathrm{Z} & 0 & -\mu \\ \mathrm{s}_\beta \mathrm{s}_\mathrm{W} m_\mathrm{Z} & -\mathrm{s}_\beta \mathrm{c}_\mathrm{W} m_\mathrm{Z} & -\mu & 0 \end{pmatrix}, \tag{2.36}$$

where $\mathrm{c}_\beta(\mathrm{s}_\beta) = \cos(\sin)\beta$ and $\mathrm{c}_\mathrm{W}(\mathrm{s}_\mathrm{W}) = \cos(\sin)\theta_\mathrm{W}$.

The mass eigenstates can be obtained by diagonalizing $\mathbf{M}_N$. This yields four neutralinos, indexed in order of increasing mass: $\widetilde{\chi}^0_1, \widetilde{\chi}^0_2$, and so on. We can see that in addition to m_Z and θ_W, the masses will be functions of only 4 MSSM parameters: M_1, M_2, μ, and β.

While in general the mass formulae are not very illuminating, in the limit $M_1, M_2, \mu \gg m_\mathrm{Z}$, the terms in the off-diagonal quadrants can be regarded as small perturbations. Then the first two neutralinos are almost pure $\widetilde{\mathrm{B}}$ and $\widetilde{\mathrm{W}}$, while the latter two are almost pure $\widetilde{\mathrm{H}}^0$. The masses are then approximately

$$m_{\widetilde{\chi}^0_1} \approx M_1 \tag{2.37}$$

$$m_{\widetilde{\chi}^0_2} \approx M_2 \tag{2.38}$$

$$m_{\widetilde{\chi}^0_3} \approx m_{\widetilde{\chi}^0_4} \approx |\mu|. \tag{2.39}$$

We now turn from the neutralino sector to the chargino sector. Again, in the gauge eigenbasis, we can denote $\psi^{\pm} = (\widetilde{\mathrm{W}}^{+}, \tilde{\mathrm{H}}_u^{+}, \widetilde{\mathrm{W}}^{-}, \tilde{\mathrm{H}}_d^{-})$. The mass matrix is given by

$$\mathbf{M}_C = \begin{pmatrix} 0 & \mathbf{X}^{\mathrm{T}} \\ \mathbf{X} & 0 \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} M_2 & gv_u \\ gv_d & \mu \end{pmatrix} = \begin{pmatrix} M_2 & \sqrt{2}\mathrm{s}_\beta m_{\mathrm{W}} \\ \sqrt{2}\mathrm{c}_\beta m_{\mathrm{W}} & \mu \end{pmatrix}. \tag{2.40}$$

We can find the squared masses of the two charginos by finding the eigenvalues of $\mathbf{X}^\dagger\mathbf{X}$. Again, we see the chargino masses only depend on a small number of MSSM parameters: M_2, μ, and β. In the limit where electroweak symmetry breaking is a small perturbation, the chargino masses are approximately

$$M_{\tilde{\chi}_1^{\pm}} \approx M_2 \tag{2.41}$$

$$M_{\tilde{\chi}_2^{\pm}} \approx |\mu|. \tag{2.42}$$

Interestingly, we see that when the effects of electroweak symmetry breaking are small, $m_{\tilde{\chi}_2^0} \approx M_{\tilde{\chi}_1^{\pm}}$. This actually holds for one more term in the expansion before the two diverge. The equivalent expression for $M_{\tilde{\chi}_2^{\pm}}$ does not hold at the next order.

2.4.3 Sfermion Masses

The masses of third-generation sfermions can be rather complicated because of mixing between left- and right-handed components. However, these mixing terms are all proportional to a power of the Yukawa coupling of the corresponding fermion. Therefore, these components are negligible for first- and second-generation sfermions. Instead, the mass terms for these sfermions come primarily from the $\widetilde{f}_L^\dagger m^2 \widetilde{f}_L$ terms in the soft SUSY-breaking term in Equation 2.27. These are unspecified by the theory, although various SUSY-breaking scenarios give more concrete predictions.

We do expect these terms to be on the order of the SUSY scale, which should be of order 1 TeV.

2.5 Experimental Consequences

Supersymmetry predicts a myriad of new particles, which are generally expected at a mass scale of around 1 TeV. Thus the signature at a collider is expected to be quite spectacular. However, SUSY has thus far not been observed at the LHC. This does not, however, rule out SUSY. In fact, only a few supersymmetric particles need be light [6]. First, the supersymmetric partner of the top quark, the stop, provides the largest correction to the Higgs boson mass. Thus, if the stop is too heavy, we are again left with an unnatural large correction to the Higgs boson mass, albeit a much smaller one than the 17 orders of magnitude we previously faced. Second, we see from Equation 2.33 that the parameter μ contributes to the electroweak symmetry breaking sector. We also saw that the masses of the Higginos were controlled by this parameter. Thus, the Higgsino masses cannot be too large, or we will get large, unnatural cancellations in the electroweak sector of the MSSM. Finally, in order to account for dark matter, the LSP must not be too heavy.

Thus, if SUSY is correct, we expect to find at least a few new particles at the TeV scale. Chapter 7 discusses a search that is sensitive to a few of these particles, the electroweak gauginos, while Chapter 6 addresses more general problems that arise in the presence of a theory that includes dark matter.

CHAPTER 3

DETECTOR

Having discussed the leading candidate for a theory of physics beyond the Standard Model, we turn now to experiment. The analyses in Chapter 6 and Chapter 7 were performed on data recorded with the Compact Muon Solenoid (CMS) detector. Here, we give an overview of how CMS the Large Hadron Collider, function.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [7] is a 27 km circumference proton-proton collider situated 100 m below ground near the CERN laboratory in Geneva, Switzerland. The data in this work were taken in collisions at 7 and 8 TeV p-p center of mass energies.

The LHC is composed of two parallel beam pipes inside a single cryostat, which contain counter-rotating proton beams. These beams cross at four different points along the LHC's circumference, each hosting one or more detectors to observe the results of the proton-proton collisions. Proton beams enter the LHC at 450 GeV from the Super Proton Synchrotron, and are accelerated to collision energies by many passes through superconducting radio-frequency cavities. The protons are steered using superconducting magnets with a peak field of 8 T.

A proton beam in the LHC is divided into roughly 10 cm long bunches. These bunches mean that proton-proton collisions are not continuous, but are divided into "bunch-crossings", which occur every 50 ns, for a crossing rate of 20 MHz. Since each bunch contains many protons, there are as many as 40 proton-proton interactions during each bunch crossing, with a mean of roughly 20 during 2012 running.

3.2 The CMS Detector

One of the detectors observing the products of the proton-proton collisions created by the LHC is the Compact Muon Solenoid (CMS) detector. The design of CMS [8] centers around a 4 T solenoid magnet, which is used to bend charged particles so that their momentum can be measured. Inside the solenoid lies a silicon tracker to measure the momenta of charged particles and electromagnetic and hadronic calorimeters to measure their energy. Outside the solenoid lies an extensive muon system, which is used to measure the momenta of muons. With this combination of detectors, CMS is capable of measuring almost all of the stable particles that are produced in the collisions. A thorough description of the CMS detector can be found in Ref. [9], but we give an overview here, focussing on the components that are used in the analyses described later.

3.2.1 Coordinate System

The CMS detector is situated at interaction point 5 of the LHC, almost directly opposite the main CERN campus. The coordinate system is defined with the origin at the nominal interaction point. The y-axis points directly upwards and the x-axis points towards the center of the LHC ring. The z-axis then points west along the beam line towards the Jura Mountains. The azimuthal angle ϕ is measured in the x-y plane, with the zero point lying along the x axis. The pseudorapidity is defined as $\eta = -\ln \tan \theta/2$, where θ is the polar angle measured from the z-axis.

3.2.2 Inner Tracking

The silicon tracker is the innermost subdetector of CMS. Its role is to precisely measure the positions and trajectories of charged particles as they leave the in-

teraction region. These measurements allow for accurate primary and secondary vertex reconstruction and, through the bending of trajectories in the magnetic field, momentum determination.

Because of the large QCD backgrounds present in a pp collider such as the LHC, leptonic decays of interesting particles are very important because of the clean signals they provide. In particular, leptonic decays of the W and Z bosons are critical to many analyses. Thus it is crucial to the physics goals of CMS that the tracker provide an accurate momentum measurement for charged leptons.

The tracker also is needed to discriminate between electrons and photons in the electromagnetic calorimeter. Additionally, the tracker must measure hadronic tracks in order to provide isolation information that allows electrons and photons to be distinguished from QCD jets. Finally, tracker momentum measurements of charged hadronic particles can be combined with information from the Hadronic Calorimeter to improve the resolution of hadronic jets.

The tracker must perform all of these functions while minimizing the amount of material that photons and electrons must pass through before reaching the electromagnetic calorimeter, as photon pair production and electron bremsstrahlung in the tracker can greatly reduce the energy resolution of the electromagnetic calorimeter [10]. The tracker must also survive the harsh radiation environment of the LHC without performance degradation, and have a fine enough spatial resolution to be able to successfully reconstruct tracks in an environment with a particle rate of approximately 10^7 Hz at the innermost point of the tracker.

Pixel Vertex Detector

The pixel vertex detector is the innermost layer of the tracker. This detector is critical for vertex reconstruction, as it provides the first measurement of a particle by the CMS detector. The pixel detector is composed of 66 million 100×150

μm^2 silicon pixels. These pixels are p-n junctions which work by collecting charges displaced in their depletion region by the passage of ionizing particles. Charge sharing between adjacent pixels allows a single-hit position resolution of 15–20 µm.

The pixel detector consists of three layers of pixels in the barrel region and two disks in the endcap region. The barrel region consists of three cylindrical layers of pixels positioned at radii of 4.4, 7.3, and 10.2 cm from the beam line and covering the region $|z| < 26.5$ cm. There are also two endcap disks at each end of the barrel region covering radii from 6 to 15 cm. The disks are located at $|z| = 34.5$ and 46.5 cm.

Silicon Strip Tracker

Outside of the pixel vertex detector lies the silicon strip tracker. This detector is composed of 1,540,000 silicon strips which collect charge when an ionizing particle passes through them. As in the pixel detector, charge sharing between adjacent strips allows single-hit position measurements with precision better than the strip pitch. The strip tracker is composed of 4 components, the tracker inner barrel (TIB), the tracker outer barrel (TOB), the tracker endcap (TEC), and the tracker inner disks (TID).

The TIB is composed of 4 concentric cylindrical layers covering the region $|z| < 65$ cm. In the TIB, the thickness of the silicon strips is 320 µm with a width that ranges from 120 to 180 µm. The strips in the first two layers of the TIB are oriented perpendicularly to each other in a "stereo" configuration in order to provide an accurate position measurement in both the $r - \phi$ and $r - z$ planes.

The TOB lies outside the TIB and consists of 6 layers covering $|z| < 110$ cm. Because of the large area to be covered, strips in the TOB are longer than those in the TIB. The less harsh radiation environment at greater distances from the

interaction point means that the silicon strips can be thicker without sustaining significant radiation damage. The strips in the TOB are 500 µm thick. The thicker strips provide a higher signal to noise ratio than those in the TIB, which is necessary to compensate for the higher capacitance of the longer strips. The first two layers of the TOB are arranged in a "stereo" configuration similar to that in the TIB.

The TEC and TID cover the high $|\eta|$ region in the tracker. The TEC is made up of 9 disks on each end of the tracker. The annular disks cover radii from 120 to 280 cm. The TID fills the gap between the TEC and the TIB with three disks on each end. Combined, they extend the tracker pseudo-rapidity coverage up to $|\eta| = 2.5$.

3.2.3 Magnet

The CMS magnet provides a 3.8 T field in which the tracking system and calorimeters sit. The magnet has a 6 m inner diameter, is 12.5 m long, and weights 220 t. The current is carried by 4 layers of NbTi windings. The magnet operates at a current of 19 kA and has a stored energy of 2.6 GJ at full field. The field of the magnet inside the tracking region is known to better than 0.1% [11].

The magnetic field outside the solenoid is controlled by a steel return yoke. This yoke is composed of 11 pieces, has a 14 m outer diammeter, and weighs 10,000 t. This leaves the muon detectors, in which the yoke is embedded, in almost the full 4 T field [9]. This field is mapped to a precision of a few percent [11].

3.2.4 Electromagnetic Calorimeter

Outside of the silicon tracker lies the electromagnetic calorimeter (ECAL). The ECAL's purpose in CMS is to measure the energies of photons and electrons as

precisely as possible, and to distinguish these particles from hadrons and QCD jets.

The ECAL is composed of 68,524 lead tungstate ($PbWO_4$) scintillating crystals. Lead tungstate was chosen because of its high density (8.3 g/cm^3) and short Molière radius (2.2 cm) and radiation length (0.89 cm). These properties allow for a very compact calorimeter. Lead tungstate crystals are also very fast, releasing 80 percent of scintillation light within 25 ns of an interaction [12].

While lead tungstate shows better resistance to radiation damage than other scintillating crystals, it does show some damage in the high-radiation environment of the LHC. During collisions, the crystals lose transparency due to the formation of color centers. The crystals naturally recover from this effect through annealing on a time scale of a few hours. Because of this, crystal transparency can vary during an LHC fill. To correct for this, the crystals are monitored by a laser system which records the changes in transparency as a function of time.

The light output of lead tungstate is low, at only about 4.5 photons per MeV of energy deposited, necessitating the use of photodetectors which provide signal multiplication. Avalanche photodiodes (APDs) were chosen for the ECAL because of their ability to function in high magnetic fields and to provide gains of up to 200.

Because of the high neutron fluence in the endcap region, APDs there would be quickly degraded by radiation damage. Thus, crystals in the endcaps are instead fitted with vacuum phototriodes (VPTs). These VPTs are single stage photomultiplier tubes with a mesh anode that allows them to operate in a magnetic field.

The ECAL is physically divided into two sections: the barrel (EB) and the endcap (EE). The barrel is cylindrical, with an inner radius of 1.29 m. The EB

is composed of 61,200 crystals, and covers pseudorapidities of $|\eta| < 1.479$. The barrel is divided into two half-barrels in z, with each half being further divided in ϕ into 18 "supermodules". Each supermodule covers 20° in ϕ. A supermodule contains 1700 crystals: 85 crystals in η by 20 in ϕ.

The ECAL endcap is positioned such that it provides the "caps" for the cylinder formed by the barrel. The endcaps are located 3.1 m from the interaction point in the z-direction and cover pseudorapidities from 1.479 to 3.0. Each endcap is divided into two halves called Dees, each of which contains 3,662 crystals. The crystals in the dees form an x-y grid, and are divided into 5 by 5 crystal arrays called supercrystals.

3.2.5 Hadronic Calorimeter

The Hadron Calorimeter (HCAL) lies directly outside the ECAL. The role of the HCAL is to measure the energy of hadrons and QCD jets. The HCAL is also used in the measurement of missing energy, so it is very important for the HCAL to have good hermeticity, covering as much as possible of the 4π solid angle.

The HCAL lies inside of the magnet coil, since the dense coil would initiate hadronic showers prematurely if particles passed through it before reaching the HCAL. Because of the space limitations created by this requirement, a major design goal of the HCAL was to fit as many interaction lengths of material as possible into the space between the ECAL and the magnet coil.

The HCAL barrel (HB) and endcap (HE) are sampling calorimeters. These detectors consist of layers of brass absorber interleaved with plastic scintillator tiles. The HCAL barrel is divided into two halves in the z-direction, covering $|\eta| < 1.3$. Each of these halves is further divided into 18 wedges, each covering 20° in ϕ. Each wedge contains 17 layers of scintillator interleaved with absorber. The

inner and outer layers of absorber are stainless steel for structural stability, while the other absorber layers are brass. Absorber layers range in thickness from 50.5 to 75 mm in the HB.

Scintillator layers are 3.7 cm thick and are sized such that they cover an area of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. This segmentation creates 32 "towers" in η. The first scintillator layer in the HB lies directly behind the ECAL electronics and cooling systems that are located in the gap between the ECAL and the HCAL. This scintillator is twice as thick as the other scintillator layers, and is intended to sample showers that develop in the material behind the ECAL.

The plastic scintillator tiles are read out with wavelength-shifting (WLS) fiber. This wavelength-shifting fiber is coupled to a clear fiber optic cable once it exits the sampling region. The clear fiber-optic cable runs to the edge of the half-barrel, where signals from each tower are optically added before being read out by a hybrid photodiode. This optical addition of signals means that the HCAL has no granularity of the radial direction. The exception to this is the two furthest towers on each end of the barrel which interlock with the HE. In these towers, layers are read out in 3 different longitudinal segments.

The HE covers a pseudorapidity range of $1.3 < |\eta| < 3.0$. It is divided into 18 sections in ϕ, matching the HB wedges. The HE contains 19 scintillator layers with the same thickness as those in the HB. The absorber layers are brass, and are 78 mm thick. This leads to a lower fraction of sampling material than in the barrel. The size of the scintillator tiles, and hence the size of the towers in η, remains the same as in the barrel for $|\eta| < 1.74$, while at higher η the size of the tiles is increased to $\Delta\phi = 0.175$ to accommodate the bending radius of the WLS fibers. The HE towers have two or three layers of segmentation in the longitudinal direction to aid in the discrimination of electromagnetic showers from hadronic

ones.

The Hadron Outer Calorimeter (HO) lies outside the magnet. The HO consists of 10 cm thick scintillator tiles, and relies primarily on the material in the magnet coil to act as an absorber. The HO serves as a final layer to sample showers that penetrate the HB and the magnet coil.

The HO is divided into five 2.5 m wide rings in z, with 12 sectors in each ring, matching the segmentation of the muon system. The center ring has two layers of scintillator at radii of 3.850 and 4.097 m with an 18 cm thick layer of iron absorber in between. The other rings use just a single layer of scintillator at a radius of 4.097 m. Readout of the HO uses the same scheme as the HB and HE.

3.2.6 Muon System

The CMS muon system makes up the outermost layer of CMS. The muon system is composed of three different types of ionization chambers. Since the muon system lies outside the HCAL and the magnet coil, muons should be the only electromagnetically interacting particles that do not shower before reaching the muon system. Thus, muon signals are generally very clean, and provide excellent discovery channels.

Drift Tubes

The muon barrel is composed primarily of drift tubes. Drift tubes were chosen for the barrel because of their excellent spatial resolution. The low neutron background rate and low muon multiplicity allow these detectors to operate with high efficiency in the barrel region, while less ideal conditions require the use of different detectors in the endcap regions.

The basic unit of the drift tube chambers is a single drift cell. One 50 µm

stainless steel anode wire runs down the middle of the cell. I-beam shaped cathodes form two walls of the cell and provide field shaping. Grounded aluminum plates form the other two walls. Extruded plastic is used to electrically insulate the cathodes from the ground plates. In operation, the anode potential is set to 3.6 kV and the cathode potential to –1.8 kV.

The gas in the drift tubes is a mixture of 80 to 90% Ar, with the remainder of the mixture being CO_2. Argon was chosen because of its consistent response to ionizing radiation, and the addition of CO_2 ensures that the electron drift velocity saturates. This saturation means that the relationship between drift time and position in the detector is linear, which allows for more accurate measurements of position.

The drift cells are arranged into superlayers, which consist of 4 planes of drift cells running in the same direction, with each plane being staggered by half of a drift cell. This forms a honeycomb lattice and minimizes dead spots. The anode wire pitch in this configuration is 4.2 cm.

Drift chambers are composed of 2 or 3 superlayers. The three innermost chambers contain 2 superlayers with drift cells running parallel to the beam line, providing position measurement in r and ϕ. Between lies a superlayer with drift cells running perpendicular to the beam line. This provides measurement in the z-direction. The outermost drift chamber contains only 2 superlayers, omitting the z-measuring superlayer. This arrangement gives a lever arm of 28 cm for momentum measurement in each chamber.

Drift chambers are arranged into 5 rings in the z-direction. The rings cover up to $|\eta| = 1.2$. Each ring is divided into 4 muon "stations" in the radial direction. The inner radii of the stations are 4.0, 4.9, 5.9, and 7.0 m. The rings consist of 12 sets of muon stations in ϕ, divided such that each drift chamber covers 30°.

The alignment of the drift chambers is staggered in the different stations, such that a straight line from the interaction point must pass through at least 3 drift chambers.

Cathode Strip Chambers

Because of the high neutron backgrounds as well as the higher muon rate, drift tubes would not work well in the endcap region. Instead, cathode strip chambers (CSC) are used. The CSCs are gaseous ionization detectors which use perpendicular planes of anode wires and cathode strips to provide good position measurement while retaining the ability to function at high rates and in high magnetic fields.

A charged particle passing through a CSC ionizes the gas in the chamber. Liberated electrons collect on the anode wires and cause an image charge to form on the cathode strips. By analyzing the charge distribution on the cathode strips in conjunction with the signal on the anode wire, a precise position measurement in all 3 dimensions can be made in a single layer.

The CSCs in the muon endcaps use 6 gas gaps. The chambers are trapezoidally shaped, with the narrower end pointing towards the beamline. The cathode strips run radially away from the beamline, while the anode wires lie approximately perpendicular to the cathodes.

Each muon endcap is divided into 4 disks in the z-direction. Each of these disks is divided into concentric rings, 3 for the innermost disk on each end and 2 for the other disks. The innermost ring of each disk contains 18 CSCs, while the other rings contain 36 CSCs apiece. CSC chambers are staggered in ϕ in order to minimize gaps in muon acceptance. The CSCs in the muon endcaps cover pseudo-rapidities up to $|\eta| = 2.4$.

Resistive Plate Chambers

In addition to the drift tubes and CSCs, resistive plate chambers (RPCs) are used in both the barrel and the endcap. RPCs are used because of their superior time resolution, which is useful for triggering and disambiguating cases with multiple hits in a single drift tube or CSC.

An RPC is a gaseous ionization detector composed of two insulating plates separated by a gas gap. Metallic strips or pads on the surface of the plates act as electrodes. The CMS RPCs use two gas gaps operated in avalanche mode, and provide good spatial resolution along with excellent time resolution. The time resolution given by the RPCs is low enough to unambiguously identify the correct bunch crossing for a muon.

In the barrel, chambers in the two innermost stations are equipped with RPCs on their inner and outer faces, while those in the outer stations are equipped with a single RPC on the inner face. In the endcaps, the first three disks have RPCs on the front face.

CHAPTER 4

EVENT RECONSTRUCTION

We have described the hardware of the CMS detector. However, we have made no mention of how electrical signals in the detector are turned into reconstructed particles. Here, we discuss the algorithms and techniques that are used to reconstruct various kinds of particles from signals in the detector.

4.1 Trigger

The first aspect of event reconstruction that must be considered is the trigger. The raw data from each event recorded by CMS takes roughly 1 MB on disk. With the LHC colliding protons at a rate of 20 MHz, writing every event to disk would require a data write speed of 20 TB/s. Furthermore, the vast majority of events at the LHC are uninteresting and well-understood di-jet events. To overcome this problem, CMS implements a two-level trigger to select only "interesting" events to be written to disk.

The first level of the trigger, called the Level 1 or L1 trigger, is implemented in hardware and firmware. The goal of the L1 trigger is to make a decision within approximately 3 µs as to whether the event should be kept, while cutting the event rate from 20 MHz to 100 kHz. It is the short decision time that drives the need to implement the L1 trigger in hardware and firmware, rather than software. Data from each bunch crossing is locally buffered on each subdetector. If the L1 trigger decides to keep an event, that data is read out and passed to the next level of trigger.

The highest layer of the L1 trigger, called the L1A, collects data from the Global Muon and Global Calorimeter Triggers. These systems, in turn, aggregate data from the various pieces of their subdetectors.

The Global Calorimeter Trigger combines information from the ECAL and HCAL. The ECAL trigger does basic electron and photon identification. This information, along with energy sums from the HCAL are passed to the Global Calorimeter Trigger, which does basic jet-finding and data aggregation such as calculating the scalar sum of energy before passing information to the L1A trigger.

The Global Muon Trigger combines data from all three muon detectors. The muon DTs, CSCs, and RPCs perform rudimentary track reconstruction in hardware and pass these tracks to the Global Muon Trigger. The Global Muon Trigger merges these tracks and calculates track quality information, which is sent to the L1A trigger.

The L1A trigger collects data from the Global Calorimeter and Muon Triggers and makes a trigger decision based upon a variety of programmable criteria. Some of these include p_{T} thresholds for various objects, jet multiplicity thresholds, or total energy thresholds. If the event is triggered on, a signal is sent to the detectors to read out their buffered data.

The second layer of triggering is performed by the High Level Trigger (HLT) [13]. This layer runs in software on a large CPU cluster. The HLT cuts the event rate from the 100 kHz L1A rate to the roughly 100 Hz that is stored permanently.

Events in the HLT are reconstructed in steps, with each step discarding events that do not meet certain criteria. For example, electrons are treated by first clustering energy in the ECAL. Requirements are made on the cluster shape and events which do not meet these requirements are discarded immediately. Next, the HLT looks for matching hits in the pixel detector, throwing out events without these hits. Finally, the full CPU-intensive track reconstruction is performed using information from the tracker. By intelligently reconstructing events in steps, CPU usage is minimized.

The exact requirements that events must meet to pass the HLT are configurable at the start of every CMS run. These requirements change regularly to reflect physics priorities and the luminosity of the LHC.

4.2 Tracks

Track reconstruction is the process of calculating charged particle trajectories from hits in the pixel detector and tracker. Track reconstruction begins with a track seed from the pixel detector, which provides an estimate of the starting track parameters. These seeds are found by looking for triplets or doublets of hits in the different pixel layers. A Kalman filter [14] is then applied iteratively to fit the track. For each layer of the tracker, the track is extrapolated to that layer using the equations of motion and the track covariance matrix. The algorithm then searches for hits in the tracker layer near the extrapolated position. This avoids the combinatorics problem which would come with searching the entire tracker for hits. Upon finding the hit, the track parameters and covariance matrix are updated using a weighted average of the extrapolated track and the measured hit.

This process is repeated until the outer layer of the tracker is reached. Then, the entire process is repeated from the outside in, starting with the location and covariance matrix at the end of the first pass. This allows the use of all hits to determine the track parameters at the inner layer, and thus the most precise extrapolation possible into the vertex region.

4.3 Electrons

Electrons in CMS are reconstructed by combining information from the ECAL and the tracker. First, clusters are created in the ECAL. These clusters are seeded by

crystals with large energy deposits and expanded to include surrounding crystals. These clusters are then combined into "superclusters" which group together clusters in a way which attempts to take into account the spread in ϕ of an electron's energy deposition due to radiation.

Once all of the superclusters are created, the electron reconstruction algorithm looks for track stubs in the pixel detector. Rather than searching the entire pixel detector, the supercluster centroids are extrapolated back to the pixels under both charge hypotheses, and only these areas are searched. Then, the track finding algorithm is applied, starting from the pixel seed. Rather than the generic Kalman filter, a specialized Gaussian Sum Filter (GSF) [15, 16] is used. This filter is a generalization of a Kalman filter, and takes into account the non-Gaussian nature of energy losses of electrons in material.

Finally, the electron's energy is determined by a resolution-weighted average of the ECAL energy and track p_{T}. In general, the ECAL dominates this combination, but the tracker starts to become important for electrons below $p_{\mathrm{T}} \sim 15$ GeV.

4.4 Muons

Compared to the reconstruction of electrons, the reconstruction of muons is fairly straightforward. Tracks are found separately in the muon system and the tracker as previously described. Next, these tracks are extrapolated to a common surface that lies between the tracker and the muon system. Any tracks that are sufficiently close to each other are combined into a global muon track. Finally, the track is re-fit using all of the hits from both the tracker and the muon system to ensure the best possible determination of the track parameters.

4.5 Particle Flow

Because of the excellent resolution of the CMS tracker, an alternative event reconstruction can be used which takes advantage of all of the information available from the various subdetectors for a global description of an event. This reconstruction method, referred to as particle-flow event reconstruction [17], outputs a list of all observed particles in an event, including charged and neutral hadrons. It can be used as an alternative reconstruction technique for electrons and muons, but its main strength lies in the improved resolution it provides for jet and missing energy measurements. It is thus the default event reconstruction used for these quantities.

The particle flow algorithm starts by collecting information from all of the subdetectors: tracks from the tracker, energy clusters from the ECAL and HCAL, and tracks from the muon system. It then iteratively links together objects that are near each other in η and ϕ. These linked objects are interpreted as particles.

The algorithm starts by linking tracks in the muon system with those in the tracker. A track which provides a good fit when tracker and muon information is combined is classified as a muon. The corresponding tracks are removed from the event, and an estimate of the energy a muon would deposit in the calorimeter is subtracted from the cells that the track crossed.

Next, the algorithm looks for tracks in the tracker which point to a cluster in the hadronic or electromagnetic calorimeter. The algorithm also looks for ECAL clusters that are tangent to the track at a point where it passed through a tracker layer. This is done in order to associate bremsstrahlung photons with a charged track. The charged tracks are classified as either electrons or charged hadrons depending on their relative energy deposition in the ECAL and HCAL. Importantly, an estimate of their energy comes from a weighted combination of tracker and calorimeter measurements. This provides a large improvement in resolution

for charged hadrons over that which would be obtained from calorimeter measurements alone, as the track resolution is much better for tracks up to p_{T} values of a few hundred GeV.

Finally, HCAL and ECAL clusters that are not associated with any tracks are categorized as neutral hadrons or photons, respectively.

4.6 Hadronic Jets

Quarks and gluons in the final state of a collision are unstable and hadronize. Thus, they appear in the detector as a spray of particles, including electrons, photons, muons, and neutral and charged hadrons. These are clustered into jets using a jet clustering algorithm.

The jet clustering algorithm takes as an input the list of individual particles from particle flow that were not determined to be isolated leptons or photons. The clustering is performed using the anti$-k_{\mathrm{T}}$ algorithm [18] with a distance parameter of $R = 0.5$.

The anti$-k_{\mathrm{T}}$ algorithm falls into the class of sequential recombination jet clustering algorithms. These algorithms work by defining a distance measure d_{ij} between any two entities and a distance measure d_{iB} between an entity and the beam. The algorithm then combines particles by sequentially adding together four-momenta of the entities with the smallest distance. If the smallest distance is d_{iB}, the entity i is labelled as a jet and removed from the collection.

In the anti$-k_{\mathrm{T}}$ algorithm, the distance measures are defined as

$$d_{ij} = \min(k_{\mathrm{T}i}^{-2}, k_{\mathrm{T}j}^{-2}) \frac{\Delta y_{ij}^2 + \Delta \phi_{ij}^2}{R^2} \tag{4.1}$$

$$d_{iB} = k_{\mathrm{T}i}^{-2} \tag{4.2}$$

where y is the rapidity, k_{T} is the transverse momentum, and R is the previously

mentioned distance parameter. This distance parameter controls the size of the jets, as it affects how soon the algorithm runs into d_{iB} as the smallest distance. Larger values of R give larger jets.

4.7 Missing Transverse Energy

The measurement of missing transverse momentum, also called missing transverse energy (MET), sometimes denoted as $\not{p}_{\mathrm{T}}$, is extremely important in searches for new physics. The concept of MET relies on the fact that the initial state particles in a collision have 0 momenta in the plane perpendicular to the beam. Thus, the final state should also have momenta that sum vectorially to 0 in the transverse direction. Deviations from 0 can be attributed to particles such as neutrinos, which carry momentum away from the collision without leaving any signal in the detector. It is also expected that any dark matter particle would behave in this way if it was produced in a collision. Therefore, MET measurements are critical to searches for models of physics that contain a dark matter candidate.

In principle, MET is straightforward to calculate. The transverse momenta of the particles returned by the particle flow algorithm are simply added vectorially. The negative of this vector sum gives the MET vector.

Despite its apparent simplicity, it is extremely challenging to gain a full understanding of the MET performance of the detector. Because MET is a summation over all particles in an event, a mismeasurement of any particle's kinematics will lead to a non-zero MET value. The MET resolution of the CMS detector has been the subject of a great deal of study [19]. Unphysical or mismeasured MET often contributes to backgrounds of searches for new physics.

CHAPTER 5

KINEMATIC ENDPOINT VARIABLES

So far, we have discussed the Standard Model of particle physics and several its problems. We then discussed one extension to the SM that might solve these problems: supersymmetry. Next, we discussed the detector and reconstruction techniques we use to study these theories. We turn now to a more specific problem.

There has been a great deal of work on searches for physics that deviates from the Standard Model. However, the next step in the process has received somewhat less attention. When and if evidence of new physics is discovered, how will we determine which theory describes what we have observed? Even within the framework of supersymmetry, there is a huge parameter space, but it may be the case that the new physics is not supersymmetry, but something else.

One of the first orders of business will be to measure properties of new particles that have been discovered, starting with their masses. One of the very problems that led us to look for new physics in the first place promises to complicate this measurement: dark matter. In order to account for the dark matter observed in astrophysics, new physics must include particles with the right properties to be dark matter. The particles must be massive, stable, and interact through only the weak and gravitational interactions.

This has several implications. First, since these new particles interact only through the weak and gravitational interactions, they will generally escape a detector like CMS without leaving behind any signal. Second, the dark matter particle must be stable. By far the easiest way to do this is to introduce a parity into a theory, under which new physics particles are odd and SM particles are even. The canonical example of this is R-parity in supersymmetry, but it also appears in other theories.

The dark matter particle must be the lightest new particle, otherwise it would not be stable. This means that, in order to conserve this new parity, any decay of a new physics particle will have a dark matter particle in the final state. The parity also has an impact on how new physics particles are produced. Because the initial state is composed of only SM particles with even parity, the total initial parity is even. The final state must also have even parity. Thus, a collision must produce an even number of new physics particles, with a minimum of two.

These factors combine to tell us something about what a new physics final state should look like. It should have an even number of dark matter particles, most likely two. This presents a problem for mass measurements, as it leaves events kinematically underconstrained due to the information that is lost with the two undetected dark matter particles. Thus, mass measurements will present new challenges and require the development of new experimental techniques. These new techniques must be independent of the underlying model, since one would hope to use the mass measurements to help determine that model. This rules out a large class of techniques that have been used to measure Standard Model particle masses in underconstrained events, for example, in dileptonic $\mathrm{t\bar{t}}$ decays.

There has been a great deal of work in the phenomenology community on this problem. See Ref [20] for a full review. Here, we discuss one class of techniques which relies on kinematic variables that exhibit endpoint structures. Since these variables rely purely on kinematics, they are independent of the underlying physics model and provide a good solution to the problem.

5.1 M_{T}

The first variable which used a kinematic endpoint to measure a particle mass in a hadron collider was the transverse mass M_{T}. The M_{T} variable was used to measure the W boson mass in leptonic decays at the UA1 and UA2 experiments [21, 22].

Consider an event with the decay $W \to \ell\nu$. Assume that the neutrino is the only source of $\not{p}_T$, that is $\mathbf{p}_T^\nu = \not{\mathbf{p}}_T$. Then M_T is defined as

$$M_T^2 = m_\nu^2 + m_\ell^2 + 2(E_T^\ell E_T^\nu - \mathbf{p}_T^\ell \cdot \mathbf{p}_T^\nu) \tag{5.1}$$

$$= m_\nu^2 + m_\ell^2 + 2(E_T^\ell \not{p}_T - \mathbf{p}_T^\ell \cdot \not{\mathbf{p}}_T). \tag{5.2}$$

When the lepton and neutrino come from a W boson decay, M_T has an upper bound. Consider the case when W boson is at rest in the lab frame. In this case, Equation 5.1 is maximized when the lepton and neutrino are in the transverse plane. Then, approximating the lepton and neutrino as massless, $\mathbf{p}_T^\nu \cdot \mathbf{p}_T^\ell = -E_T^\nu E_T^\ell$ and $E_T^\nu = E_T^\ell = m_W/2$. This gives a maximum value $M_T^{\max} = m_W$. Since M_T is by construction invariant under boosts in the transverse plane, this upper endpoint must hold even when the W boson is not at rest in the lab frame. Thus,

$$m_W \geq M_T. \tag{5.3}$$

The knowledge of the maximum value of M_T for a given W mass gives a way to relate the observable quantities to the mass of the W boson; by looking for the observed maximum of the M_T distribution, we obtain the mass of the W boson.

5.2 M_{T2}

We now make the problem harder by considering pair-produced particles with decays similar to those of the W boson. As we discussed, this is a natural event topology to consider in new physics theories which provide a dark matter candidate. The topology gives us events that look like the one shown in Figure 5.1. The particles labeled P (for parent) represent the pair-produced particles, and those labelled C (for child) represent the dark matter candidate particles,which are not detected. The particles labelled v represent SM particles (jets, leptons, *etc.*), which are measured by the detector.

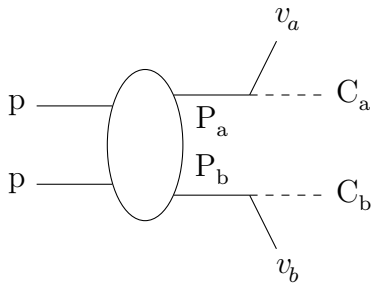


Figure 5.1: The generic event topology in which we would like to measure particle masses. Two protons collide. Through some unknown process, two particles P are pair-produced. Each decays to a visible particle v and an undetected particle C.

With two unobserved particles in the final state, the event kinematics are even less constrainted than in $W \to \ell\nu$ decays. There, the $\not{p}_T$ came from only the neutrino, so we could assign $\mathbf{p}_T^{\nu} = \not{\mathbf{p}}_T$. Here, the $\not{p}_T$ comes from two particles, leaving us with only the constraint $\mathbf{p}_T^{C_a} + \mathbf{p}_T^{C_b} = \not{\mathbf{p}}_T$.

5.2.1 Definition

Let us proceed as if we did in fact know $\mathbf{p}_T^{C_a}$ and $\mathbf{p}_T^{C_b}$ in order to devise a way to generalize M_T to events with two identical decay chains. We start by calculating two M_T variables, one for each decay chain. Call them M_{Ta} and M_{Tb}. We know from our discussion of M_T that $m_P \geq M_{Ta}$ and $m_P \geq M_{Tb}$. It then follows that

$$m_P \geq \max(M_{Ta}, M_{Tb}). \tag{5.4}$$

Now consider our lack of knowledge of $\mathbf{p}_T^{C_a}$ and $\mathbf{p}_T^{C_b}$. We know that Equation 5.4 holds for the true values of these vectors. If we proceed by minimizing the quantity on the right hand side of Equation 5.4 over the allowed configurations of $\mathbf{p}_T^{C_a}$ and $\mathbf{p}_T^{C_b}$, that is, those that obey the constraint $\mathbf{p}_T^{C_a} + \mathbf{p}_T^{C_b} = \not{\mathbf{p}}_T$, we are guaranteed to obtain a quantity that is less than or equal to the value that we would obtain with

the true values of $\mathbf{p}_{\mathrm{T}}^{\mathrm{C_a}}$ and $\mathbf{p}_{\mathrm{T}}^{\mathrm{C_b}}$. Thus,

$$m_{\mathrm{P}} \geq \min_{\mathbf{p}_{\mathrm{T}}^{\mathrm{C_a}}+\mathbf{p}_{\mathrm{T}}^{\mathrm{C_b}}=\not{\mathbf{p}}_{\mathrm{T}}} \left[\max\left(M_{\mathrm{T}a}, M_{\mathrm{T}b}\right)\right] \tag{5.5}$$

Finally, there is one more complication: we do not know the mass of the C particles, which is needed to calculate M_{T}. For the moment, we can simply make everything a function of this unknown mass. We denote our guess for this unknown mass m_χ. We have now defined the M_{T2} variable, first proposed in Ref. [23].

$$M_{\mathrm{T2}}(m_\chi) \equiv \min_{\mathbf{p}_{\mathrm{T}}^{\mathrm{C_a}}+\mathbf{p}_{\mathrm{T}}^{\mathrm{C_b}}=\not{\mathbf{p}}_{\mathrm{T}}} \left[\max\left(M_{\mathrm{T}a}(m_\chi), M_{\mathrm{T}b}(m_\chi)\right)\right] \tag{5.6}$$

Note that M_{T2} is really a function of m_χ, rather than a variable.

M_{T2} is constructed so that

$$m_{\mathrm{P}} \geq M_{\mathrm{T2}}(m_\chi = m_{\mathrm{C}}). \tag{5.7}$$

While it is not immediately obvious that equality can be achieved in Equation 5.7, it has been shown [24] that this is indeed the case. Thus, were it not for the unknown test mass, M_{T2} could be used in a fashion similar to M_{T}: one measures the upper endpoint, and thus m_{P}. Working around this unknown test mass has been the subject of a great deal of theoretical work, which will be described shortly.

5.2.2 Interpretation

At this point, it is perhaps useful to introduce an alternative physical interpretation of M_{T2}, first described by Cheng and Han in Ref. [25]. Consider an event of the form shown in Figure 5.1. We substitute again for m_{C} a test mass m_χ. Making very few assumptions, we can write down a set of constraints for an event that follow from simple relativistic kinematics:

$$p_{\mathrm{C_a}}^2 = m_\chi^2 \tag{5.8}$$

$$p_{\mathrm{C_b}}^2 = m_\chi^2 \tag{5.9}$$

$$(p_{\mathrm{C_a}} + p_{v_a})^2 = m_{\mathrm{P}}^2 \tag{5.10}$$

$$(p_{\mathrm{C_b}} + p_{v_b})^2 = m_{\mathrm{P}}^2 \tag{5.11}$$

$$p_{\mathrm{C_a}}^x + p_{\mathrm{C_b}}^x = \not{p}^x \tag{5.12}$$

$$p_{\mathrm{C_a}}^y + p_{\mathrm{C_b}}^y = \not{p}^y. \tag{5.13}$$

While there are not enough equations to solve for all of the unknowns, the requirements that all physical quantities be real and that energies be positive place some restrictions on several of the quantities. Cheng and Han show that for a given test mass, a physical solution exists only if $m_{\mathrm{P}} \geq M_{\mathrm{T2}}(m_\chi)$. Thus, we can view M_{T2} as the boundary of the region allowed by the constraints.

This alternative interpretation of M_{T2} is useful for two reasons. First, it provides another way of thinking about M_{T2} that can be useful when studying and understanding its properties. Second, it led to an algorithm, developed by Cheng and Han, that calculates M_{T2} values an order of magnitude faster than the numerical minimization used in the simplest implementation.

5.2.3 Properties

A great deal of theoretical work has been done on exploring the properties the M_{T2} variable. Here, we summarize the important ones. In particular, we summarize the M_{T2} "kink", which has received a great deal of theoretical attention. As neither of the analyses in this work rely on the kink, we focus on describing the underlying properties of the M_{T2} variable which lead to the kink.

Dependence on test mass

We return now to the question of the unknown test mass. If the visible particles v_a and v_b are massless, the dependence of the endpoint on the test mass is quite simple [26, 27]:

$$M_{\mathrm{T2}}^{\max}(m_\chi) = \mu + \sqrt{\mu^2 + m_\chi^2}, \tag{5.14}$$

where

$$\mu = \frac{m_{\mathrm{P}}^2 - m_{\mathrm{C}}^2}{2m_{\mathrm{P}}}. \tag{5.15}$$

Thus, a test mass can simply be *chosen*. The combination of masses μ can then be extracted independent of the choice of test mass.

Behavior with upstream pt

So far, we have considered the case where the PP system is at rest in the transverse lab frame. This is seldom the case in a hadron collider such as the LHC because of the presence of initial state radiation (ISR). Consider Figure 5.2, which is our familiar decay chain, but with some upstream particles that the PP system is recoiling against. These upstream particles can originate in two ways: first, through ISR, and second, when the PP system is just one step in a longer decay chain.

In general, any non-zero upstream momentum will change the M_{T2} endpoint, making Equation 5.7 invalid. However, there is an analytical formula for the dependence of the endpoint on the magnitude of the upstream momentum U_{T} [28]:

$$M_{\mathrm{T2}}^{\max}(m_\chi, U_{\mathrm{T}}) = \begin{cases} F_L(m_\chi, U_{\mathrm{T}}) & \text{for } m_\chi \leq m_{\mathrm{C}} \\ F_R(m_\chi, U_{\mathrm{T}}) & \text{for } m_\chi \geq m_{\mathrm{C}}. \end{cases} \tag{5.16}$$

The exact forms of F_L and F_R are complicated and not particularly illuminating, so we will not reproduce them here. They are shown graphically in Figure 5.3. Note that while the function is piecewise continuous, there is a discontinuitiy in

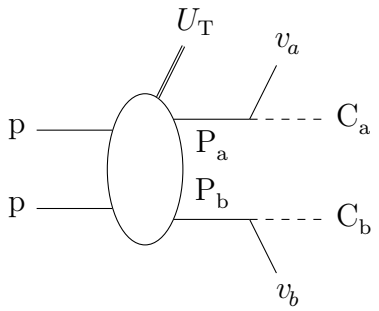


Figure 5.2: The generic event topology in which we would like to measure particle masses. Two protons collide. Through some unknown process, two particles P are either pair-produced or appear as products of upstream decays. They recoil against some upstream objects with total transverse momentum U_{T}. Each P particle decays to a visible particle v and an undetected particle C.

the slope at $m_\chi = m_{\mathrm{C}}$. It is important to note that the dependence on U_{T} vanishes when $m_\chi = m_{\mathrm{C}}$. Thus, if the child mass is known, the test mass can be set to the child mass and this issue vanishes.

Kink

If the mass of the child particle is unknown, one can try to take advantage of the discontinuity in slope shown for $U_{\mathrm{T}} > 0$ in Figure 5.3. The basic idea is to measure $M_{\mathrm{T2}}^{\mathrm{max}}$ as a function of the test mass m_χ. When the function is plotted, there will be a "kink" when $m_\chi = m_{\mathrm{C}}$, allowing one to determine the correct value of m_{C}. The value of $M_{\mathrm{T2}}^{\mathrm{max}}$ at this point will then give m_{P}.

The application of this method has been simulated both at the parton-level [29] and with a fast detector simulation [30], but the utility of this method for data recorded with a realistic detector has not been thoroughly explored. In particular, neither of these works addressed the difficulty of extracting the location of the kink in the $M_{\mathrm{T2}}^{\mathrm{max}}$ vs m_χ plot when the data points have uncertainties and are strongly correlated with each other.

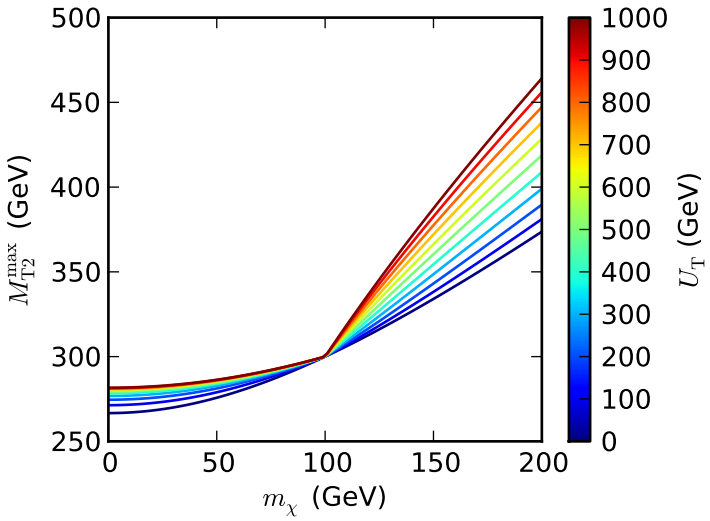


Figure 5.3: Plots of Equation 5.16 for $m_P = 300$ GeV, $m_C = 100$ GeV with different values of U_T. The magnitude of U_T is indicated by the color of the line. For non-zero U_T, a kink appears when $m_\chi = m_C$, and the value of M_{T2}^{max} at this point gives m_P.

Dependence on visible particle mass

There is another way that the kink can appear: when the visible particles have masses which vary event-to-event. This can occur, for example when the P particles decay via a 3-body decay or a multistep decay chain, as in Figure 5.4. In this case, the visible particle four-momenta are actually the sum of the four-momenta of two particles: $p_{v_a} = p_{v_{a1}} + p_{v_{a2}}$ and $p_{v_b} = p_{v_{b1}} + p_{v_{b2}}$.

This type of kink was first observed in Ref. [31] and was further studied in Ref. [29]. It was shown that for $U_T = 0$,

$$M_{T2}^{max}(m_\chi) = \begin{cases} F(m_\chi, m_<) & \text{for } m_\chi \leq m_C \\ F(m_\chi, m_>) & \text{for } m_\chi \geq m_C, \end{cases} \tag{5.17}$$

where $m_<$ and $m_>$ are the minimum and maximum kinematically allowed visible particle masses, respectively. As before, the exact form of F is not particularly interesting, but we do note several important properties. First, as before, the two

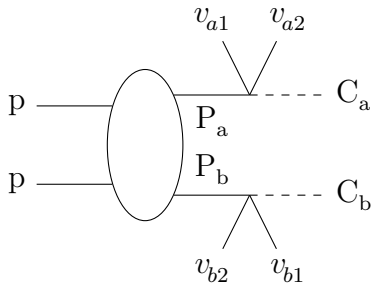


Figure 5.4: The event topology with three-body decays. Two protons collide. Through some unknown process, two particles P are pair-produced. Each decays via a three-body decay to a pair of visible particles v_1 and v_2 and an undetected particle C.

terms are equal at $m_\chi = m_C$, but the derivatives are discontinuous, giving rise to a kink at this point. Also note that, by construction, when the visible particles have constant mass, $m_< = m_> = m_v$, the two sides of Equation 5.17 are equal, and the kink disappears.

Because this kink was discovered first, it has been studied in simulation somewhat more thoroughly than the kink created by $U_T > 0$ [27–29, 31]. Nevertheless, it remains to be seen how accurately the location of the kink can be extracted from data recorded in a real detector.

5.3 M_{CT}

A simpler alternative to the M_{T2} variable is provided by the M_{CT} variable [32]. This variable was invented to solve the same problem as M_{T2}: mass measurement in decay topologies like that in Figure 5.1. It also has many of the same properties as M_{T2}, but has the benefit of being much simpler to work with.

5.3.1 Definition

M_{CT} is defined in terms of only the visible particles v_a and v_b.

$$M_{\mathrm{CT}}^2 = m^2(v_a) + m_(^2 v_b) + 2\left[E_{\mathrm{T}}(v_a)E_{\mathrm{T}}(v_b) + \mathbf{p}_{\mathrm{T}}(v_a)\cdot\mathbf{p}_{\mathrm{T}}(v_b)\right] \tag{5.18}$$

M_{CT} has an upper endpoint similar to that of M_{T2}

$$M_{\mathrm{CT}}^{\mathrm{max}} = \frac{m_{\mathrm{P}}^2 - m_{\mathrm{C}}^2}{m_{\mathrm{P}}} \tag{5.19}$$

when $m_{v_a} = m_{v_b} = 0$. Ref. [32] shows that this upper bound is indeed saturated when the visible particles are co-linear. Thus, the M_{CT} endpoint can be used in a similar fashion to M_{T2}, but with fewer complications from test masses and $\not{p}_{\mathrm{T}}$ measurements.

5.3.2 Interpretation

It is somewhat difficult to arrive at a physical interpretation for M_{CT}. The motivation for its introduction in Ref. [32] was to find a function of the visible particle masses that was invariant under equal magnitude back-to-back boosts of the visible particles. Such boosts are expected, as the visible particles are decay products of heavier particles. This was achieved by modifying the definition of the transverse mass given in Equation 5.1 to introduce a negative sign in front of the scalar product of the momenta. Thus, a variable which was invariant under co-linear boosts was made invariant under contra-linear boosts.

Also helpful in interpreting the M_{CT} variable is the observation [33] that $M_{\mathrm{CT}} = M_{\mathrm{T2}}$ under certain assumptions, namely $m_\chi = 0$, $m_{v_a} = m_{v_b} = 0$, and $U_{\mathrm{T}} = 0$. We also note that the M_{CT} endpoint is simply twice the M_{T2} mass parameter, $M_{\mathrm{CT}}^{\mathrm{max}} = 2\mu$, using μ from Equation 5.15.

5.3.3 Properties

Many properties of the M_{CT} endpoint were explored in Refs. [34, 35]. While the behavior of M_{CT} is not as rich as that of M_{T2}, we nevertheless catalogue some if its properties here for for the sake of completeness.

Equation 5.19 holds only for $m_{v_a} = m_{v_b} = 0$. For non-zero visible particle masses, the endpoint becomes

$$M_{\mathrm{CT}}^{2\,\mathrm{max}} = m^2(v_a) + m^2(v_b) + 2m_{v_a}m_{v_b}\cosh[\zeta(v_a) + \zeta(v_b)], \tag{5.20}$$

where

$$\sinh\zeta(v) = \frac{\lambda^{\frac{1}{2}}(m^2(\mathrm{P}), m^2(\mathrm{C}), m^2(v))}{2m_{\mathrm{P}}m_v}, \tag{5.21}$$

and

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz. \tag{5.22}$$

Like the M_{T2} endpoint, the M_{CT} endpoint depends on U_{T}. However, unlike M_{T2}, there is no point where the dependence of $M_{\mathrm{CT}}^{\mathrm{max}}$ on U_{T} vanishes. The dependence is given in the notation of [35] by

$$M_{\mathrm{CT}}^{2\,\mathrm{max}} = m^2(v_a) + m^2(v_b) + 2m_{v_a}m_{v_b}\cosh[2\eta + \zeta(v_a) + \zeta(v_b)], \tag{5.23}$$

where

$$\sinh\eta = \frac{U_{\mathrm{T}}}{2m_{\mathrm{P}}}, \qquad \cosh\eta = \sqrt{1 + \left(\frac{U_{\mathrm{T}}}{2m_{\mathrm{P}}}\right)^2}. \tag{5.24}$$

5.4 Subsystems

One of the major drawbacks of the kinematic variables discussed here is that their endpoints depend on a combination of unknown masses. Thus, if we ignore the M_{T2} kink, a single endpoint measurement cannot fully determine both unknown masses in events of the type shown in Figure 5.1.

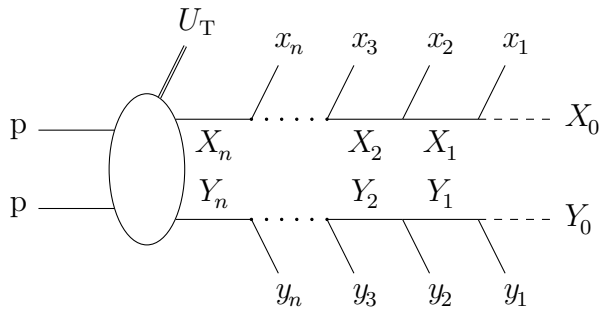


Figure 5.5: A generic n-step symmetric pair of decays. The incoming protons interact through some unspecified process, producing identical particles X_n and Y_n and possibly ISR with momentum summing to U_{T}. The two produced particles decay through symmetric multistep decays. The final state has $2n$ visible particles, plus unobserved particles X_0 and Y_0.

We can expand the problem slightly by considering multistep decays like the one shown in Figure 5.5. The event can then be divided into several "subsystems", each with an endpoint that can be measured and related to the unknown masses. For two-step or longer decay chains, a full determination of all the unknown masses can be made.

This idea was first explored in Refs. [32, 33] and was more clearly defined and simulated in Ref. [28]. Considering a multistep decay chain like that in Figure 5.5, we use the notation of Ref. [28] to label different subsystems.

Steps in the decay chain are indexed as shown in Figure 5.5. Each subsystem is labelled by 3 indices, npc. The first index n denotes the length of the decay chain, defined as the index of particle furthest up the decay chain. Thus, for a given decay chain, n is always the same. We then specify the parent and child particles, which give the indices p and c, respectively. These are the particles are the analogues of P and C in Figure 5.1. All visible particles emitted between the parent and child particles are used as the measured visible particles. In the case where there are several visible particles emitted from each decay chain, particles from the same chain are combined by summing their four-momenta, and the kinematic variables

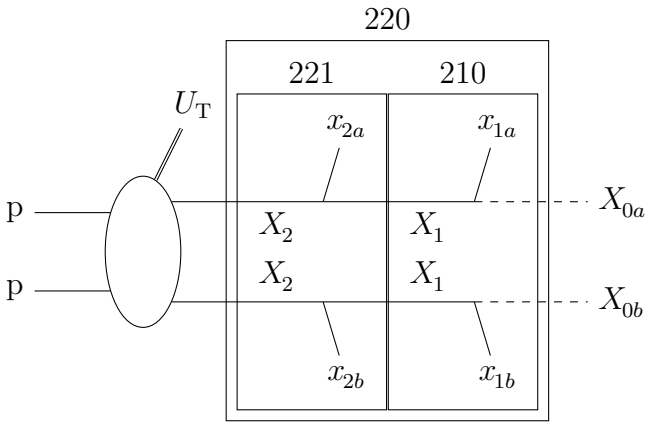


Figure 5.6: The three different subsystems for an $n = 2$ step decay. The different subsystems treat different particles as parents, children, and visible particles.

are calculated as they would be for the event in Figure 5.4. Visible particles below the child particle are added to the $\not{\mathbf{p}}_{\mathrm{T}}$ vector to form the downstream p_{T} vector, which we will denote $\mathbf{D}_{\mathrm{T}}$. In the non-subsystem cases we have discussed previously, we had $\mathbf{D}_{\mathrm{T}} = \not{\mathbf{p}}_{\mathrm{T}}$.

Figure 5.6 shows the three possible subsystems that can be defined for an $n = 2$ step decay chain. The 221 subsystem treats the X_2 particles as the parents, the X_1 particles as the children, and the x_2 particles as visible. The M_{T2} variable would be calculated as

$$M_{\mathrm{T2}}^{221}(m_\chi) \equiv \min_{\mathbf{p}_{\mathrm{T}}(X_{1a})+\mathbf{p}_{\mathrm{T}}(X_{1b})=\mathbf{D}_{\mathrm{T}}^{221}} \left[\max\left(M_{\mathrm{T}a}^{221}(m_\chi), M_{\mathrm{T}b}^{221}(m_\chi)\right)\right], \tag{5.25}$$

where $\mathbf{D}_{\mathrm{T}}^{221} = \mathbf{p}_{\mathrm{T}}(x_{2a}) + \mathbf{p}_{\mathrm{T}}(x_{2b}) + \not{\mathbf{p}}_{\mathrm{T}}$,

$$M_{\mathrm{T}}^{221} = \sqrt{m_\chi^2 + m_{x_2}^2 + 2(E_{\mathrm{T}}^{X_1} E_{\mathrm{T}}^{x_2} - \mathbf{p}_{\mathrm{T}}^{X_1} \cdot \mathbf{p}_{\mathrm{T}}^{x_2})}, \tag{5.26}$$

and m_χ is the test mass for X_1. An analogous formula applies for M_{CT}. The mass parameter for the 221 system gives the mass combination

$$\mu^{221} = \frac{m^2(X_2) - m^2(X_1)}{m(X_2)}. \tag{5.27}$$

The 210 subsystem treats the X_1 particles as the parents, the X_0 particles as

the children, and the x_1 particles as visible. We then have

$$M_{\mathrm{T2}}^{210}(m_\chi) \equiv \min_{\mathbf{p}_{\mathrm{T}}(X_{0a})+\mathbf{p}_{\mathrm{T}}(X_{0b})=\mathbf{D}_{\mathrm{T}}^{210}} \left[\max\left(M_{\mathrm{T}a}^{210}(m_\chi), M_{\mathrm{T}b}^{210}(m_\chi)\right)\right], \tag{5.28}$$

where $\mathbf{D}_{\mathrm{T}}^{210} = \not{\mathbf{p}}_{\mathrm{T}}$,

$$M_{\mathrm{T}}^{210} = \sqrt{m_\chi^2 + m_{x_1}^2 + 2(E_{\mathrm{T}}^{X_0} E_{\mathrm{T}}^{x_1} - \mathbf{p}_{\mathrm{T}}^{X_0} \cdot \mathbf{p}_{\mathrm{T}}^{x_1})}, \tag{5.29}$$

and m_χ is the test mass for X_0. The mass parameter is

$$\mu^{210} = \frac{m^2(X_1) - m^2(X_0)}{m(X_1)}. \tag{5.30}$$

Finally, we come to the 220 subsystem. In this case, we treat the X_2 particles as parents and the X_0 particles as children. The visible particles are composites formed by adding together x_2 and x_1. For the moment, we ignore the combinatorial problem of how to determine which visible particles share the same decay chains. M_{T2} is then given by

$$M_{\mathrm{T2}}^{220}(m_\chi) \equiv \min_{\mathbf{p}_{\mathrm{T}}(X_{0a})+\mathbf{p}_{\mathrm{T}}(X_{0b})=\mathbf{D}_{\mathrm{T}}^{220}} \left[\max\left(M_{\mathrm{T}a}^{220}(m_\chi), M_{\mathrm{T}b}^{220}(m_\chi)\right)\right], \tag{5.31}$$

where $\mathbf{D}_{\mathrm{T}}^{220} = \not{\mathbf{p}}_{\mathrm{T}}$ and

$$M_{\mathrm{T}}^{220} = \sqrt{m_\chi^2 + m_{x_1+x_2}^2 + 2(E_{\mathrm{T}}^{X_0} E_{\mathrm{T}}^{x_1+x_2} - \mathbf{p}_{\mathrm{T}}^{X_0} \cdot \mathbf{p}_{\mathrm{T}}^{x_1+x_2})}, \tag{5.32}$$

and m_χ is the test mass for X_0. The mass parameter is gives

$$\mu^{220} = \frac{m^2(X_2) - m^2(X_0)}{m(X_2)}. \tag{5.33}$$

Given Equation 5.27, Equation 5.30, and Equation 5.33, we have three independent equations for three unknown parameters. Thus, we can determine the masses of X_2, X_1, and X_0. We can use either M_{T2} or M_{CT} to determine the μ parameters, which gives the freedom to choose whichever works better for a particular situation.

It is worth noting at this point that there is a fourth endpoint that can be measured in this decay: that of the x_2-x_1 invariant mass. This can provide useful supplementary information, or can replace the measurement of μ^{220}. The endpoint is given by

$$m^{\max}_{x_1-x_2} = \frac{1}{m(X_1)}\sqrt{(m^2(X_2)-m^2(X_1))(m^2(X_1)-m^2(X_0)}. \tag{5.34}$$

5.5 Projection

We have shown that both the M_{T2} and M_{CT} endpoints depend on the amount of upstream momentum U_{T} in each event. Even in the absence of ISR, some subsystem variables have contributions to U_{T} from non-ISR process. For example, in the 210 subsystem, the x_2 particles contribute to U_{T}. This dependence must be accounted for.

Since we have formulas describing the dependence of each variable on U_{T}, one could imagine approaches such as binning events by U_{T} or performing an unbinned likelihood analysis which takes into account this dependence. However, these approaches would rely upon accurate measurement of U_{T}. Since U_{T} primarily consists of low-p_{T} ISR jets and unclustered energy, which are generally poorly measured, the precision that can be attained by such attempts may be severely limited.

Instead, there have been several proposals to modify the M_{T2} and M_{CT} variables to remove the U_{T} dependence of their endpoints. First, as we have already discussed, in cases where we know the child particle mass, the U_{T} dependence vanishes from M_{T2} by setting the test mass to the child mass. However, let us consider cases where we do not know any of the masses.

It turns out that a simple modification can be applied that removes the U_{T} dependence from both M_{T2} and M_{CT}. The effects of the upstream momentum are treated as a simple boost of the PP system. Thus, all decay products also receive

this same boost. To remove the effects of the boost, we can simply consider only the components of particle momenta *perpendicular* to this boost. M_{T2} and M_{CT} are then calculated in the usual way, but using the projected momenta [30, 36].

5.5.1 $M_{\mathrm{T2}\perp}$

It is worth exploring the properties of these projected variables in a bit of detail, as the simplifications that arise from the projection process allow their properties to be explored somewhat more thoroughly than their unprojected counterparts.

The $M_{\mathrm{T2}\perp}$ variable is the projected version of M_{T2}, and is defined as

$$M_{\mathrm{T2}\perp}(m_\chi) \equiv \min_{\substack{\mathbf{p}_{\mathrm{T}\perp}^{\mathrm{C_a}} + \mathbf{p}_{\mathrm{T}\perp}^{\mathrm{C_b}} = \not{\mathbf{p}}_{\mathrm{T}\perp}}} \left[\max\left(M_{\mathrm{T}\perp a}(m_\chi), M_{\mathrm{T}\perp b}(m_\chi)\right)\right]. \tag{5.35}$$

The properties of $M_{\mathrm{T2}\perp}$ are explored in Ref. [36]. After projection, $M_{\mathrm{T2}\perp}$ can be calculated analytically when the visible particles are massless.

$$M_{\mathrm{T2}\perp} = \sqrt{A_{\mathrm{T}\perp}} + \sqrt{A_{\mathrm{T}\perp} + m_\chi^2}, \tag{5.36}$$

where

$$A_{\mathrm{T}\perp} = \frac{1}{2}\left(|\mathbf{p}_{\mathrm{T}\perp}(v_a)||\mathbf{p}_{\mathrm{T}\perp}(v_b)| + \mathbf{p}_{\mathrm{T}\perp}(v_a)\cdot\mathbf{p}_{\mathrm{T}\perp}(v_b)\right). \tag{5.37}$$

This can be re-written in terms of the angles ϕ_a and ϕ_b between the visible particles and the $\mathbf{U}_{\mathrm{T}}$ vector.

$$A_{\mathrm{T}\perp} = \frac{1}{2}|p_{\mathrm{T}}(v_a)||p_{\mathrm{T}}(v_b)|\left(|\sin\phi_a \sin\phi_b| + \sin\phi_a \sin\phi_b\right) \tag{5.38}$$

This form allows us to easily understand several of the properties of the $M_{\mathrm{T2}\perp}$ variable. First, we can immediately see that when the signs of ϕ_a and ϕ_b are opposite, $A_{\mathrm{T}\perp} = 0$ and $M_{\mathrm{T2}\perp} = m_\chi$. Figure 5.7 shows the situation in which this occurs. If we divide the transverse plane in half along the direction of the upstream momentum, we get $M_{\mathrm{T2}\perp} = m_\chi$ when the visible particles are in opposite halves.

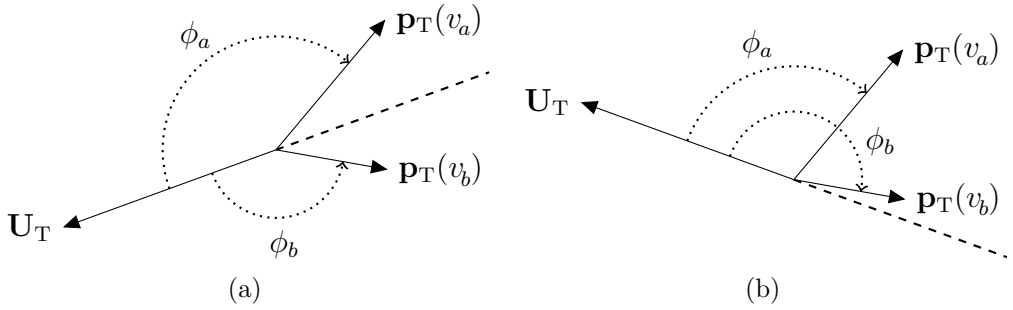


Figure 5.7: Kinematic configurations in the transverse plane where (a) ϕ_a and ϕ_b have opposite sign, and $A_{\mathrm{T}\perp} = 0$ and (b) ϕ_a and ϕ_b have the same sign, and $A_{\mathrm{T}\perp} > 0$. The dashed line divides the transverse plane in half along the direction of the transverse momentum. In (a), the two visible particles are on opposite sides of the line, and in (b), the two visible particles are on the same side of the line.

Thus, a significant fraction of events are expected to have this minimal value of $M_{\mathrm{T2}\perp}$.

Another interesting property of $M_{\mathrm{T2}\perp}$ is that an analytic form for the shape of the distribution can be derived. It is given by

$$\frac{dN}{dM_{\mathrm{T2}\perp}} = N\delta(M_{\mathrm{T2}\perp} - m_\chi) + (1-N)\frac{M_{\mathrm{T2}\perp}^4 - m_\chi^4}{\mu^2 M_{\mathrm{T2}\perp}^3} \ln\left(\frac{2\mu M_{\mathrm{T2}\perp}}{M_{\mathrm{T2}\perp}^2 - m_\chi^2}\right). \quad (5.39)$$

Here, N is the fraction of events for which $M_{\mathrm{CT}\perp} = m_\chi$, as discussed above and δ is the Dirac delta function indicating that they all have the same value. μ is the mass parameter defined in Equation 5.15.

5.5.2 $M_{\mathrm{CT}\perp}$

The properties of the $M_{\mathrm{CT}\perp}$ variable are presented in Ref. [35]. The $M_{\mathrm{CT}\perp}$ variable is defined as

$$M_{\mathrm{CT}\perp}^2 = m^2(v_a) + m^2(v_b) + 2\left(E_{\mathrm{T}\perp}(v_a)E_{\mathrm{T}\perp}(v_b) + \mathbf{p}_{\mathrm{T}\perp}(v_a)\cdot\mathbf{p}_{\mathrm{T}\perp}(v_b)\right), \quad (5.40)$$

where

$$E_{\mathrm{T}\perp} = \sqrt{p_{\mathrm{T}\perp}^2 + m^2}. \quad (5.41)$$

This can again be simplified in terms of the angles between the visible particles and the upstream momentum, ϕ_a and ϕ_b. For massless visible particles,

$$M^2_{\mathrm{CT}\perp} = 2p_{\mathrm{T}}(v_a)p_{\mathrm{T}}(v_b)\left(|\sin\phi_a \sin\phi_b| + \sin\phi_a \sin\phi_b\right). \tag{5.42}$$

We see that when the visible particles are massless, $M_{\mathrm{CT}\perp} = M_{\mathrm{T2}\perp}(m_\chi = 0)$. This is expected, as we already know that $M_{\mathrm{CT}} = M_{\mathrm{T2}}(m_\chi = 0)$ when the visible particles are massless and $U_{\mathrm{T}} = 0$. By projecting against $\mathbf{U}_{\mathrm{T}}$, we have accomplished the same thing.

Given this correspondence with $M_{\mathrm{T2}\perp}$, it is easy to derive the distribution of $M_{\mathrm{CT}\perp}$:

$$\frac{dN}{dM_{\mathrm{CT}}} = N\delta(M_{\mathrm{CT}\perp}) + (1-N)\frac{M_{\mathrm{CT}\perp}}{4\mu^2}\ln\left(\frac{2\mu}{M_{\mathrm{CT}\perp}}\right). \tag{5.43}$$

5.6 Previous Applications of Kinematic Endpoint Variables

There have been several studies which make use of kinematic endpoint variables in either real collider data or with realistic detector simulations and backgrounds. The CDF Collaboration has used the M_{T2} variable to measure the top quark mass in dileptonic decays [37, 38]. They use a template method to extract the mass, with templates derived from Monte Carlo simulation, and thus do not make use of the model-independent nature of kinematic variables.

Two studies by the LHC/LC Study Group [39] show prospects for measuring slepton and squark masses using M_{T2}. The determination of slepton mass again relies on simulated events, while the squark mass study does a fit to the M_{T2} distribution to find the endpoint. This study, however, does not address the issue of the unknown child mass, and measures only the squark-LSP mass difference.

Finally, several studies have been done which use the M_{T2} variable in searches rather than for mass measurement. The CMS Collaboration has done a search for

supersymmetry in all-hadronic events using M_{T2} as a discriminating variable [40]. The ATLAS Collaboration has conducted searches for sbottom quarks decaying to $\mathrm{b}\widetilde{\chi}_1^0$ and for chargino and slepton pair production using the M_{CT} variable [41, 42]. All of these searches take advantage of the endpoints of these distributions by looking for new physics in events which fall above the endpoints.

CHAPTER 6

MASS MEASUREMENT USING KINEMATIC ENDPOINTS

If kinematic variables are to be used to study new physics, it is important to get an idea of their performance and limitations when applied to real data. In service of this goal, this chapter presents work on using kinematic endpoint variables to measure masses in dileptonic $t\bar{t}$ events. With two neutrinos in the final state, these events resemble new physics events with two dark matter particles in the final state. Application of kinematic endpoint methods to dileptonic $t\bar{t}$ decays has been proposed several times in the literature as a way to test their efficacy using currently available data [28, 34, 43].

In order to take into account the limitations one would face when applying the method to new physics, care is taken in this analysis to avoid relying too much on the fact that decays of the top quark are very well understood. For this reason, although simulation was used to design and test the analysis method, the final result does not rely directly on simulation.

When information about the masses of the other particles in the events is added, the top-quark mass can be measured with precision comparable to other methods. Because of the limited reliance by this analysis on Monte Carlo simulation, it provides a measurement which is complimentary to traditional methods [44–46].

6.1 Event Selection

In this analysis, events are reconstructed using the particle flow algorithm described in Section 4.5. We want to select events in which the process $t\bar{t} \to bW^+\bar{b}W^- \to b\ell^+\bar{\nu}\bar{b}\ell^-\nu$ occurs. The strategy for achieving this is to require events to have two well-isolated electrons or muons, missing transverse momentum, and two jets that are tagged as coming from b quarks.

Each lepton must pass certain requirements in order to be considered "good". Electrons are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. The η requirement ensures that we select only electrons within the tracker acceptance. We also make several requirements designed to discriminate against jets and photons faking electrons. We require electrons to be isolated. We define the relative isolation of an electron to be the sum p_{T} of all particle flow objects within $\Delta R = 0.3$ (where $\Delta R^2 = \Delta\phi^2 + \Delta\eta^2$) divided by the electron p_{T}. This quantity is required to be less than 0.2 for the electron to be considered isolated. This isolation requirement helps to reject electrons in jets, which almost always have other tracks nearby.

We place similar requirements on muons. A muon must have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$. The η requirement keeps muons within the muon trigger acceptance. We also apply a series of requirements on the reconstructed muon track in order to ensure selection of a well-measured muon. As with electrons, we require muons to have a relative isolation less than 0.2 in order to reject muons in jets.

We require events to have an opposite-sign ee, $\mu\mu$, or eμ pair in which both leptons pass the above quality cuts. In events with a same-flavor pair, that is, ee or $\mu\mu$ pairs, the dilepton invariant mass must be at least 15 GeV away from the Z boson mass at 91 GeV.

Next, we require events with a same-flavor dilepton pair to have $\not{p}_{\mathrm{T}} > 30$ GeV. This further reduces backgrounds from events with single Z boson production and unphysical $\not{p}_{\mathrm{T}}$ from jet mismeasurements. In eμ events, smaller Z backgrounds allow us to lower the $\not{p}_{\mathrm{T}}$ requirement to 20 GeV.

Finally, we require events to have two jets that are identified as likely to have originated from b quarks. b-quark jet candidates are required to have $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$, where again the η requirement keeps them within the acceptance of the tracker. Because of the dependence of kinematic endpoint variables on the

masses of visible particles, we require the jet mass, defined as the invariant mass of the jet constituents, to be less than 40 GeV.

The likelihood of a jet originating from a b quark is evaluated using the Combined Secondary Vertex b-tagger, described in Ref. [47]. This b-tagger takes advantage of the displaced secondary vertex often seen in b-quark jets (hereafter "b jets"), using a multivariate approach which combines information from the charged tracks in the jet with information about secondary vertices. The tagger produces a discriminator value for each jet, with high discriminator values indicating jets that are likely to be b jets. This analysis uses the "loose" working point described in Ref. [47], which requires a b-tagging discriminator > 0.244. This gives a b-jet identification efficiency of 85%, while giving a non-b jet only a roughly 10% chance of being incorrectly tagged as a b jet.

This analysis is performed on 5.0 fb^{-1} of data recorded at a pp center-of-mass energy of 7 TeV. Table 6.1 shows the number of events predicted by simulation to pass the selection criteria, divided by process. Here and throughout the chapter, $\text{t}\bar{\text{t}}$ events are simulated using `MC@NLO 3.41` [48], diboson samples using `PYTHIA 6.4` [49], and all other processes using `MadGraph 5.1.1.0` [50]. The interaction of particles with the CMS detector is modeled using a `GEANT4`-based [51] simulation [8].

Table 6.1: Event counts predicted by simulation for signal and background processes

Process	Count
$\text{t}\bar{\text{t}}$ signal (no τ)	7000
$\text{t}\bar{\text{t}}$ ($\tau \to \ell\nu\nu$)	1100
Single top	270
Z/γ^*	77
Non-dilepton $\text{t}\bar{\text{t}}$	55
Diboson	14
W+jets	9

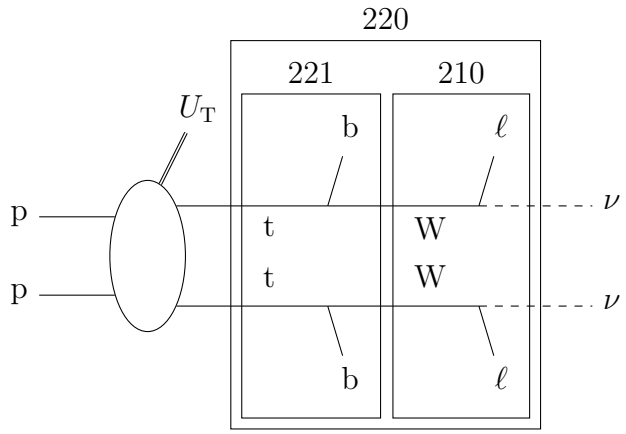


Figure 6.1: The subsystems described in Section 5.4 for dileptonic $\mathrm{t\bar{t}}$ decays.

6.2 Analysis Strategy

The goal of this analysis is to use kinematic endpoint variables to extract the top-quark, W boson, and neutrino masses from dileptonic top-quark pair decays, $\mathrm{t\bar{t}} \rightarrow \mathrm{bW^+\bar{b}W^-} \rightarrow \mathrm{b}\ell^+\overline{\nu}\mathrm{\bar{b}}\ell^-\nu$. This is a two-step decay chain with three "unknown" masses. To extract these masses, we combine the subsystem concept described in Section 5.4 with the projected M_{T2} variable, $M_{\mathrm{T2}\perp}$, described in Subsection 5.5.1.

The subsystems for these decays are shown in Figure 6.1. The 221 subsystem treats the top quarks as the parent particles, the W bosons as the children and the b jets as the visible particles. The 210 subsystem treats the W bosons as the parent particles, the neutrinos as the child particles, and the leptons as the visible particles. The 220 subsystem is not used in this analysis. We instead get our third constraint from the b-ℓ invariant mass, as described in Section 5.4.

We choose to use the $M_{\mathrm{T2}\perp}$ variable because it has several nice properties. The upstream momentum independence provided by the projection procedure is important, especially in the 210 subsystem, which always has two b jets contributing to $\mathbf{U}_{\mathrm{T}}$, even in the absence of ISR. This $\mathbf{U}_{\mathrm{T}}$-independence could be attained using vanilla M_{T2} if we were willing to assume the correct test mass, but this would

defeat the purpose of considering all three measured masses to be "unknown".

Taking the b-jets to be massless, we can write the endpoints of the three variables in terms of the masses to be measured.

$$\max(M_{\mathrm{T2}\perp}^{221})(m_\chi) = \mu + \sqrt{\mu^2 + m_\chi^2}, \qquad \mu = \frac{m_\mathrm{t}^2 - m_\mathrm{W}^2}{m_\mathrm{t}} \tag{6.1}$$

$$\max(M_{\mathrm{T2}\perp}^{210})(m_\chi) = \mu + \sqrt{\mu^2 + m_\chi^2}, \qquad \mu = \frac{m_\mathrm{W}^2 - m_\nu^2}{m_\mathrm{W}} \tag{6.2}$$

$$\max(m_{\mathrm{b}\ell}) = \frac{\sqrt{(m_\mathrm{t}^2 - m_\mathrm{W}^2)(m_\mathrm{W}^2 - m_\nu^2)}}{m_\mathrm{W}} \tag{6.3}$$

Note that we do need to specify a test mass m_χ for each of the $M_{\mathrm{T2}\perp}$ variables, but since the dependence of the endpoint on the test mass is known, we may pick any test mass we like. Simulations show that the results do not depend on the chosen value of the test mass. For simplicity, we choose $m_\chi = 0$ GeV for the 210 subsystem and $m_\chi = m_\mathrm{W} = 80.4$ GeV for the 221 subsystem.

Conceptually, we should measure the three endpoints on the left-hand sides of these equations, then invert the equations to solve for the masses. In practice, it is easiest to parameterize the endpoints in terms of the masses in order to avoid regions where the equations may have no real solutions. Thus, the fit maximizes the likelihood as a function of the masses directly. In the case of the neutrino mass, we fit for the m_ν^2, since this is the only way that the neutrino mass appears in the endpoint equations.

We now come to the problem of how to fit the endpoints. We use an unbinned maximum-likelihood method, so the problem lies in specifying the likelihood as a function of the parameters. First, we factorize the likelihood by assuming that the three distributions are uncorrelated, an assumption supported by studies in simulation. (As an aside, this is the major reason that $M_{\mathrm{T2}\perp}^{220}$ is not used: for many

events it is equal to $m_{\mathrm{b}\ell}$). The likelihood function is then of the form

$$\mathcal{L}(\mathbf{x}|\mathbf{m},\alpha) = \mathcal{L}^{221}(\mathbf{x}^{221}|\mathbf{m}) \cdot \mathcal{L}^{210}(\mathbf{x}^{210}|\mathbf{m}) \cdot \mathcal{L}^{\mathrm{b}\ell}(\mathbf{x}^{\mathrm{b}\ell}|\mathbf{m}), \tag{6.4}$$

where $\mathbf{x}$ represents the observed data, $\mathbf{m} = (m_{\mathrm{t}}, m_{\mathrm{W}}, m_{\nu})$ the unknown masses, and α any nuisance fit parameters.

We model the likelihood for each distribution using three terms: the underlying "true" distribution of the data, the response function of the detector, and a background function which describes the distribution of non-signal events. Thus, the likelihood for a single measurement x looks like

$$\mathcal{L}(x|\mathbf{m}) = \alpha f(x|\mathbf{m}) \circ R(x) + (1-\alpha)B(x), \tag{6.5}$$

where $f(x|\mathbf{m})$ is the true distribution, $R(x)$ is the detector response function, $\circ$ represents convolution, $B(x)$ is the background distribution, and α is the probability that an event is signal.

Let us now step through each piece of this likelihood function.

6.3 Fit Procedure

6.3.1 Signal Shapes

In the case of the $M_{\mathrm{T2}\perp}$ variables, we know the true shapes of the underlying distributions, which are given in Equation 5.39. In practice, however, these shapes are distorted in complicated ways by the various kinematic requirements in the event selection criteria. Thus we take the slightly more pragmatic approach of approximating the signal shape as linear in a small region near the endpoint. The function used to describe this linear approximation is

$$f(x|x_{\max}(\mathbf{m})) = \begin{cases} \mathcal{N}(x_{\max} - x) & \text{for } x \leq x_{\max} \\ 0 & \text{for } x > x_{\max} \end{cases}, \tag{6.6}$$

where $\mathcal{N}$ is fixed by normalization and $x_{\max}$ is the endpoint of the distribution (which is a function of the unknown masses).

Note that this treatment introduces several complications. Since the function is certainly not linear across its entire range, we must choose a range in which this approximation holds. While any function can be approximated as linear on a small enough interval, decreasing the size of the interval decreases the number of events that are used to extract the endpoint. In the end, we choose a fit range that minimizes the dependence of the result on the exact choice of the fit range. We assign a systematic uncertainty due to this choice. Evaluation of this uncertainty is discussed in detail in Subsection 6.6.2.

6.3.2 Backgrounds

Combinatorial

When calculating $m_{\mathrm{b}\ell}$ values, we do not have any way of determining which leptons and b jets come from the decays of same top quarks. There are two different ways to pair the b jets with the leptons. If we calculate $m_{\mathrm{b}\ell}$ using a b jet and lepton from different top-quark decay chains, there is no reason for it to respect the upper bound given by Equation 6.3. Thus, there is, in principle, a background from events where we choose the wrong pairing.

Since we are already using an approximation for the shape of the distribution near the endpoint, we do not necessarily need to eliminate all backgrounds; we simply need to ensure that they do not produce events past the endpoint. If we can do this, it is not even necessary to use the background term $B(x)$. For this combinatorial background, the simplest solution would be to simply select the pairing where the larger $m_{\mathrm{b}\ell}$ value is smallest. However, this frequently discards the correct pairing, and we can do slightly better. To do this, we use the following

strategy.

Let A and a denote the two bℓ invariant masses calculated from one of the two possible $b\ell$ pair assignments, and let B and b denote the masses calculated from the other assignment. Choose the labeling of masses such that $A > a$ and $B > b$. Without making any assumptions about which pairing is correct, one can order the masses from largest to smallest; there are six possible orderings. For example the ordering A, a, B, b means that the first assignment (Aa) has masses which are both larger than the masses in the second assignment (Bb). Table 6.2 shows the six possibilities. For each mass ordering shown in the left column, the right column shows the mass values that will be selected for use in the $m_{\mathrm{b}\ell}$ fit. For any given event only one row of the table applies.

It is perhaps easiest to understand the algorithm by going through a few lines of the table. In the first row, the values are ordered $AaBb$, and we always select Bb. There are two possibilities: If Bb is the correct pairing, we have selected the correct pairing. On the other hand, if Aa is the correct pairing, we have selected the wrong pairing. However, we know that $A < m_{\mathrm{b}\ell}^{\mathrm{max}}$ and $B, b < A$, so we are guaranteed to have selected values below $m_{\mathrm{b}\ell}^{\mathrm{max}}$. This corresponds to the simple approach of selecting the pairing where the larger $m_{\mathrm{b}\ell}$ is smallest.

The second line of the table is the same as the first, with the As and Bs swapped. The third row proves more interesting. The ordering is $ABab$. Imagine Aa are the values from the correct pairing. We select B, a, b, giving two values from the incorrect pairing and one from the correct pairing. However, $A < m_{\mathrm{b}\ell}^{\mathrm{max}}$ and $B, a, b < A$, so we know that all three values we picked must be below the endpoint. On the other hand, if Bb are the correct values and we pick B, a, b, we have selected two values from the correct pairing and one from the incorrect pairing. Again, since $B < m_{\mathrm{b}\ell}^{\mathrm{max}}$ and $a < B$, we are guaranteed that the value we

Table 6.2: $m_{\mathrm{b}\ell}$ Orderings

Ordering	Selection
AaBb	B,b
BbAa	A,a
ABab	B,a,b
BAab	A,a,b
ABba	B,b,a
BAba	A,b,a

chose from the incorrect pairing is less than the endpoint. The remaining entries in the table are permutations of this example.

This selection algorithm ensures that all $m_{\mathrm{b}\ell}$ values used in the fits will not exceed the endpoint because of an incorrect b-ℓ pairing, while keeping the maximal amount of information.

$\mathrm{t\bar{t}}$ with τ s

One fundamentally irreducible background that must be considered comes from dileptonic $\mathrm{t\bar{t}}$ decays where one or both of the leptons is a τ, which then decays leptonically via $\tau \to \ell\nu_\ell\nu_\tau$. These events do not match the assumed event topology of Figure 6.1, but are very similar to our signal events, in which both of the leptons are electrons or muons. Thus, we might be concerned that these events could spoil our endpoint measurements.

However, we are saved again by the fact that we do not need to worry about backgrounds that cannot extend past the endpoint. It is fairly straightforward to see that this is the case for $\mathrm{t\bar{t}}$ decays with τs. When a τ decays leptonically, some fraction of its energy and momentum is carried away by the neutrinos. Thus, the energy of the lepton decay product, which is used to calculate $m_{\mathrm{b}\ell}$ and $M^{210}_{\mathrm{T2}\perp}$, is always less than that of the τ. Since $M_{\mathrm{T2}\perp}$ and $m_{\mathrm{b}\ell}$ both decrease monotonically

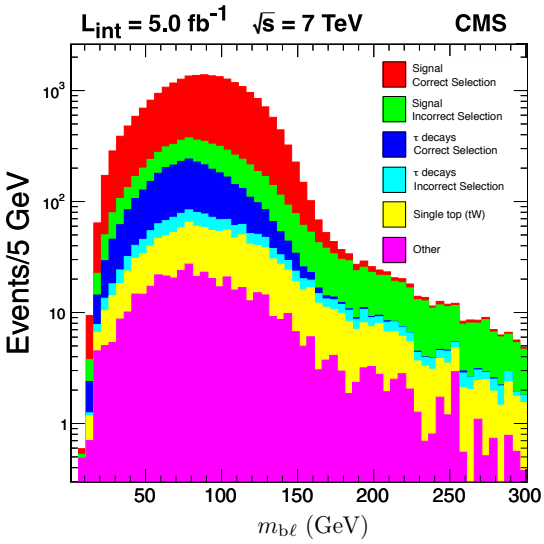


Figure 6.2: Composition of simulated event samples. Note that the events which are signal events, but where an incorrect jet was selected, contribute most of the background above the endpoint.

with decreasing lepton momenta, neither value can be greater than the value that would have been calculated with the full τ momentum, which must be less than or equal to the endpoint.

Mis-Tag

Finally, we come to a background that *does* cause events beyond the endpoint. This background arises when one or both of the selected b jets is not, in fact, a b jet, but a light jet that the b-tagging algorithm mis-tagged. These events are, for the most part, still dileptonic $\mathrm{t\bar{t}}$ decays. However, for these events, we calculate $M_{\mathrm{T2}\perp}^{221}$ and $m_{\mathrm{b}\ell}$ using a jet that is not a product of the top quark decay. Thus, these values can go above the endpoint. A shown in Figure 6.2, mis-tag events account for almost all of the events that are significantly above the endpoint in simulation.

This mis-tag background is essentially unavoidable, since there is no such thing as a perfect b-tagging algorithm. Therefore, we derive a data-driven model describing the shape of this background and include it in the fit via the $B(x)$ term

in Equation 6.18. The amount of background is left to float in the fit.

We start by identifying a sample of events that have the same characteristics as the mis-tag background. In this case, we are looking for dileptonic $t\bar{t}$ events in which we calculate kinematic variables using one jet that comes from a top-quark decay and one jet that does not. By doing this, we neglect the background that arises when both selected jets are mis-tagged. Simulations show that the effects of ignoring this are small. Our desired control sample is fairly straightforward to achieve: we intentionally select one jet that is unlikely to be a b jet by requiring it to have a low b-tagging discriminator.

Thus, the event selection for this control sample is the same as for the signal, except one of the jets is required to have a b-tagging discriminator < 0.05 instead of the usual > 0.244. This selection gives a sample that is about 70% events of the desired type. The remaining events are mostly events where we select the two jets from the top quarks, even when we're trying not to. Studies show that the presence of these signal events in our control sample does not affect the final result of the fit.

Having obtained a control sample that models our background, we now must derive functions $B(x)$ describing the distributions of $M_{\mathrm{T2}\perp}^{221}$ and $m_{\mathrm{b}\ell}$ for these control samples. To do this, we use an adaptive kernel density estimator (AKDE) [52]. Given a sample from a distribution, this method gives a non-parametric estimate of the probability density function (PDF) from which that sample was drawn. The PDF is estimated by a normalized sum of kernels, with one kernel per observation. We use a Gaussian kernel. Thus, the background function is given by

$$B(x) = \mathcal{N} \sum_{y_i} \frac{1}{b} \exp\left(-\frac{(x-y_i)^2}{2b^2}\right), \tag{6.7}$$

where y_i is the value of the variable in question for the ith event in the control sample. The value of $\mathcal{N}$ is set by normalization, but the parameter b, called the

bandwidth, must be specified. The bandwidth can be thought of as a smoothing parameter; larger values give smoother functions. There are a variety of heuristics for choosing the value of b based on the size and variance of a data sample [52, 53]. In general, a smaller bandwidth should be chosen for a sample where observations are more densely clustered.

In this case, we are interested in a region where the density of events is changing rapidly. Thus, the bandwidth which may be appropriate for one region may not be appropriate for another region. To overcome this, we use an adaptive bandwidth. We first perform a "pilot estimate" of $B(x)$ using a fixed bandwidth. Call this pilot estimate $B_0(x)$. We then do a second pass, where the bandwidth is different for each y_i. The bandwidth is given by

$$b_i = h_0 \left(\frac{B_0(y_i)}{g} \right)^{-\frac{1}{2}}, \tag{6.8}$$

where g is the geometric mean of $B_0(y_i)$ evaluated over all the y_i, and serves as a normalization factor. This adaptive scheme scales the bandwidth with the local density of observations, so that less dense regions have larger bandwidths. We must still choose an overall bandwidth scale factor h_0. We set this to 5 GeV, but we find that the final result is not sensitive to the choice of this value.

Figure 6.3 shows the performance of this background shape estimate in simulation. We see that the function produced by the AKDE of the control sample follows closely the true shape of the background.

6.3.3 Detector Resolution

The response of the detector plays an important role in the shape of the kinematic distributions near the endpoint. We model the response using a function that describes the probability of measuring a particular value of the variable given the

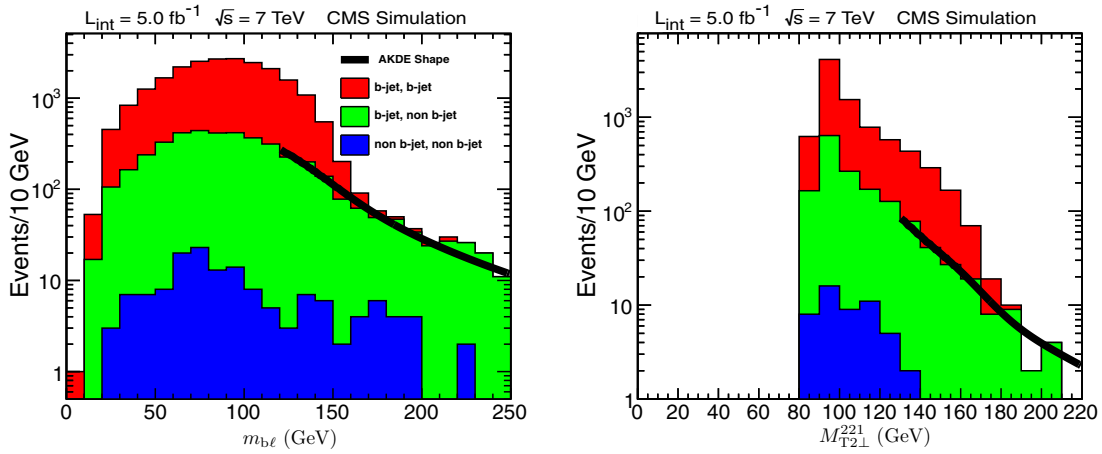


Figure 6.3: The background shape predicted by the AKDE from the tag anti-tag control sample plotted over the signal and background in simulation. We see that the predicted shape matches the true shape of the background very well.

true value. The detector response enters the likelihood through the term $R(x)$ in Equation 6.18. Since this term is convoluted with the true distribution of the variable, $f(x|\mathbf{m})$, it has the effect of "smearing" the true distribution.

Typically, uncertainties are propagated using a linear approximation. That is, if we have some derived quantity $f(x, y)$, and we assume that x and y have Gaussian uncertainties with width σ_x, σ_y, then f also has Gaussian uncertainty with width

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2. \tag{6.9}$$

This, however, assumes that the dependence of f on x and y is well-described by its first derivatives. If the derivatives of f change substantially over the length scales σ_x or σ_y, this is not the case.

For the $m_{\mathrm{b}\ell}$ variable, the derivative with respect to the observables is smooth almost everywhere, and this approach works well. We take into account the resolutions in p_{T}, η, and ϕ for both jets. We neglect the lepton resolutions, as they are much smaller than the jet resolutions.

When we consider the resolutions of the $M_{\mathrm{T2}\perp}$ variables, we find that linear

error propagation from the underlying object kinematics does not work. Intuitively, we can see that this is a consequence of the projection process. Consider the case of massless visible particles, where A_{T}, which goes into the $M_{\mathrm{T2}\perp}$ calculation, is given by Equation 5.38. Let us consider what happens if ϕ_a is mismeasured. We note that when ϕ_a crosses zero, it moves A_{T} from a region where it is always zero to a region where it depends on the values of the angles. At $\phi_a = 0$, the derivative of A_{T} is discontinuous. Thus, the assumptions underlying linear error propagation do not hold.

Let us consider the general case. For the following discussion, we denote the variable being considered (e.g. $M_{\mathrm{T2}\perp}^{210}$) as x and the vector of observed kinematic quantities as $\mathbf{k}$. Calculating the kinematic variable is a matter of calculating $x(\mathbf{k})$. Note that, in general, the function $x(\mathbf{k})$ is many-to-one; specifying $\mathbf{k}$ specifies x, but the inverse is not true.

We wish to calculate the probability density function $P(x_m|x_t)$, where x_m and x_t are measured and true values of the kinematic variable, respectively. To begin, we use Bayes rule to invert the relationship

$$P(x_m|x_t) = P(x_t|x_m)\frac{P(x_m)}{P(x_t)} \approx P(x_t|x_m), \tag{6.10}$$

where we have approximated the ratio of prior probabilities as 1. We can now phrase the problem as: "Given our measured value of x, what might the true value have been?". In order to make the problem more tractable, we slightly change it to calculate $P(x_t|\mathbf{k}_m)$: "Given our measured value of $\mathbf{k}$, what might the true value of x been?" By doing this, we are implicitly ignoring contributions from other kinematic configurations that might give the same x_m value as the measured configuration.

210

The case of $M_{\mathrm{T2}\perp}^{210}$ is somewhat special. Here, the leptons are much more precisely measured than the upstream momentum angle, so we can take them to have perfect resolution. In this case, the upstream momentum angle is the only component of $\mathbf{k}$, so $\mathbf{k} = \Phi$. In this special case, the function $x(k)$ is locally invertible, and we can take advantage of that fact.

We denote the angle of $\mathbf{U}_{\mathrm{T}}$ as Φ. It is convenient to measure this angle such that the midline between the two lepton $\mathbf{p}_{\mathrm{T}}$ vectors corresponds to $\Phi = 0$. Then, $\phi_a = -\phi_b \equiv \phi$. Setting the test mass to 0, Equation 5.38 becomes

$$x(\Phi) = \begin{cases} \sqrt{4p_{\mathrm{T}1}p_{\mathrm{T}2}\sin(\Phi-\phi)\sin(\Phi+\phi)} & \text{for } |\Phi| > |\phi| \\ 0 & \text{otherwise} \end{cases}. \tag{6.11}$$

For the non-zero portion of x, the equation is locally invertible:

$$\Phi(x) = \frac{1}{2}\cos^{-1}\left(2p_{\mathrm{T}1}p_{\mathrm{T}2} - \frac{x^2}{2p_{\mathrm{T}1}p_{\mathrm{T}2}} - 1\right). \tag{6.12}$$

Note that this equation is actually multi-valued. Even if we restrict ourselves to $0 < \Phi < 2\pi$, there are in general two values of Φ for a given x.

We can now express $P(x_t|\Phi_m)$ in terms of $P(\Phi_t|\Phi_m)$.

$$P(x_t|\Phi_m) = \sum_{\Phi'} \left|\frac{d\Phi}{dx}\right|_{\Phi=\Phi'} P(\Phi'|\Phi_m), \tag{6.13}$$

where Φ' are solutions to $x(\Phi) = x_t$, which can be easily found from Equation 6.12. There are in general two solutions, which we will denote Φ_{t1} and Φ_{t2}. The derivative $d\Phi/dx$ can be found using Equation 6.12. The calculation is messy, but straightforward.

The PDF $P(\Phi_t|\Phi_m)$ is related to the Φ resolution function, $P(\Phi_m|\Phi_t)$, by the ratio of the probabilities of Φ_m and Φ_t (from Bayes rule). We set this factor to 1, as

our events should be distributed evenly in Φ. The resolution function is modeled by a normal distribution, with width equal to the uncertainty on Φ_m:

$$P(\Phi_t|\Phi_m) = \text{Normal}(\Phi_t - \Phi_m, \sigma_\Phi) \tag{6.14}$$

Putting the pieces together, we are left with our resolution function

$$R(x) = \text{Normal}(\Phi_{t1} - \Phi_m, \sigma_\Phi) \left|\frac{d\Phi}{dx}\right|_{\Phi_{t1}} + \text{Normal}(\Phi_{t2} - \Phi_m, \sigma_\Phi) \left|\frac{d\Phi}{dx}\right|_{\Phi_{t2}}. \tag{6.15}$$

221

In the case of $M_{\text{T2}\perp}^{221}$, the visible particles are jets. Because jet resolutions are comparable to the upstream momentum angle resolution, we cannot assume that any of the resolutions are small. Thus, the vector of kinematics $\mathbf{k}$ is composed of p_T, η, and ϕ for each jet, plus the upstream angle Φ. In this case, there are not just two values of $\mathbf{k}$ that give the same value of x, but many. Thus, the approach used for the 210 subsystem will not work here.

Instead, we use a numerical approach to evaluating $P(x_t|\mathbf{k}_m)$. The basic idea is to use our knowledge of the resolutions to generate a Monte Carlo sample of true kinematics $\mathbf{k}_t$ that might have lead to the measured value $\mathbf{k}_m$. We then calculate $x_t = x(\mathbf{k}_t)$ for each Monte Carlo event. This approach bypasses the need for calculating $d\mathbf{k}/dx$, as phase-space effects are accounted for automatically.

We wish to generate a sample from $P(\mathbf{k}_t|\mathbf{k}_m)$, when the quantity we know is $P(\mathbf{k}_m|\mathbf{k}_t)$. We can calculate one PDF in terms of the other using Bayes' rule:

$$P(\mathbf{k}_t|\mathbf{k}_m) = P(\mathbf{k}_m|\mathbf{k}_t)\frac{P(\mathbf{k}_t)}{P(\mathbf{k}_m)} \tag{6.16}$$

In the 210 case, we had $\mathbf{k} = \Phi$. We claimed that $P(\Phi)$ was flat, and so we could ignore the ratio of prior probabilities. In the 221 case, the same is true for η and ϕ, and we ignore the ratio of their probabilities. However, $P(p_\text{T})$ shows a strong decrease with increasing p_T. Therefore, we must account for the ratio of prior

probabilities using the formula. To do this, we parameterize the jet p_{T} distribution for $\mathrm{t\bar{t}}$ events from simulation, then use this parameterization to evaluate the ratio in Equation 6.16. The final result is very insensitive to the choice of jet p_{T} parameterization, as long as something reasonable is used.

To generate a sample from $P(\mathbf{k}_t|\mathbf{k}_m)$, we first generate a sample of true kinematics from $P(\mathbf{k}_m|\mathbf{k}_t) = \mathrm{Normal}(\mathbf{k}_m - \mathbf{k}_t, \sigma)$, which is a multivariate normal distribution with variances given by the jet resolutions. We then assign each sampled event a weight given by $P(p_{\mathrm{T}t1}) \cdot P(p_{\mathrm{T}t2})/P(p_{\mathrm{T}m1}) \cdot P(p_{\mathrm{T}m2})$. The histogram of x_t values calculated from the sample, with these weights applied, gives us an approximate sample from the distribution $P(x_t|\mathbf{k}_m)$. This, in turn, approximates $P(x_m|x_t)$.

To perform the convolution with the signal shape, we assume that the dependence on the true value is of the form $P(x_m|x_t) = f(x_m - x_t)$. This makes the numerical convolution similar to the one which would be done when using linear error propagation: smearing the true function with a fixed-shape resolution function. The convolution is of the form

$$\int f(x_t)g(x_m - x_t)dx_t, \tag{6.17}$$

where $g(x_m - x_t)$ is the shape we generated numerically.

6.3.4 Uncertainty Estimation

Since we are fitting masses using a maximum likelihood estimator, the usual approach to estimating the uncertainty on the result would be to derive the uncertainty from the Fisher information matrix or to find the point in parameter space where the log-likelihood decreases by 0.5. However, in this case, there are two problems with these approaches. First, these approaches only give the correct result in

the asymptotic limit of infinitely large samples [54]. Second, they only account for sources of variance that are expressed in the likelihood. In this analysis, there is an additional source of variance: the control sample used in the AKDE estimate of the mis-tag background. The likelihood only contains information about the shape of the AKDE, not its uncertainty.

Thus, we turn to an alternative method of uncertainty estimation: bootstrapping [55]. The goal of bootstrapping is to provide an estimate of the probability density function (PDF) of a statistic. The statistic is a function of the data, and so if we knew the distribution of the data variables, we could derive, in principle, the distribution of the statistic. Even if this calculation were intractable, we could approximate the distribution of the statistic by drawing many samples from the data distribution and calculating the statistic on each sample.

When we do not know the distribution from which the data are drawn, we can approximate this distribution. In the absence of any other information, the best estimator for the data distribution is the data itself. We model the data PDF as a normalized sum of delta functions, with each observed data point contributing one delta function. We can then draw random samples from this distribution and calculate the statistic on each sample, giving an approximation to the distribution of the statistic.

The bootstrapping procedure gives an alternative way to think about this process. Rather than sampling from the sum of many delta functions, we simply sample from the observed data, *with replacement*. Sampling with replacement means that each event can be sampled more than once. Thus, if we sample N events with replacement from a dataset of size N, some events will appear multiple times, and some events will not appear at all. Because we sample with replacement, this procedure is equivalent to sampling from the delta-function PDF.

Let us formally define the bootstrapping algorithm. Consider a data sample of size N. Each data point j has observables X_i, and we wish to find the properties of the statistic $\Theta(X_{ij})$.

Repeat $n_{samples}$ **times**
  Repeat N **times**
    Choose a random number j between 0 and N;
    Add the jth data point to this sample;
  end
  Calculate Θ on this sample;
end

For large values of n_{samples}, the values of Θ generated this way approximate a sample drawn from true distribution of Θ. Thus, we can make inferences about Θ, such as estimates of the mean and variance.

In this analysis, the statistic in question is the fitted mass. We would like to find the mean and standard deviation of this statistic. Applying the bootstrapping procedure to the problem is fairly straightforward. We draw many bootstrap samples from the data by sampling from the data with replacement. Importantly, we also use a bootstrap sample of the events in the tag anti-tag control region to create the AKDE. Thus, each bootstrap sample has its own estimate of the background distribution $B(x)$. For each bootstrap sample, we calculate the fitted mass using a maximum likelihood fit. Finally, we extract the mean and standard deviation from the bootstrap distribution of the mass. The mean gives our best estimate for the mass, while the standard deviation gives the standard error on this estimate.

6.3.5 Efficiency

Another important factor that must be considered is the efficiency of selecting jets and leptons as a function of their kinematics. If this efficiency varies as a function

of any of the kinematic variables, it could lead to a bias in the fitted mass.

To correct for this effect, we take advantage of the bootstrapping procedure. When creating bootstrap samples, instead of sampling uniformly, we sample events with probability inversely proportional to their efficiency. Thus, an event in a region with low efficiency is more likely to be sampled than one in a region with high efficiency. This produces bootstrap samples in which all kinematic regions are sampled equally. This method of weighted bootstrapping has previously been used in labor force surveys to account regional variations in non-response rates [56], a problem which nicely parallels the one here, where we have kinematic variations in efficiency.

6.3.6 Summary

Since there are quite a few pieces to the analysis, we now review how the top mass is extracted. First, we draw efficiency-weighted bootstrap samples from the observed data. For each bootstrap sample, we do a maximum likelihood fit to find the masses.

The likelihood function, given by Equation 6.4, is the product of three likelihoods: one each for the $M_{\mathrm{T2}\perp}^{210}$, $M_{\mathrm{T2}\perp}^{221}$ and $m_{\mathrm{b}\ell}$ distributions. The three likelihoods are products of individual event likelihoods, which are given by Equation 6.18.

$$\mathcal{L}(x|\mathbf{m}) = \alpha f(x|\mathbf{m}) \circ R(x) + (1-\alpha)B(x), \tag{6.18}$$

The function $f(x|\mathbf{m})$ represents the true shape of the distribution in question. For all three variables, we take this to be a straight line decreasing into the endpoint and zero above the endpoint.

The function $R(x)$ is convoluted with the true shape, and accounts for imperfect detector resolution. This term varies event-by-event, depending on the event

kinematics. Calculating $R(x)$ is handled differently for each of the three variables. For $m_{\mathrm{b}\ell}$, the uncertainty is simply propagated from the observed objects using the usual linear approximation. For $M_{\mathrm{T2}\perp}^{210}$, we propagate the uncertainty using an analytical calculation, which does not make the linear approximation. The resolution on the $M_{\mathrm{T2}\perp}^{221}$ variable is calculated numerically.

The final term, $B(x)$, accounts for backgrounds which may give events past the kinematic endpoint. For the $M_{\mathrm{T2}\perp}^{210}$ variable, there is no such background, so $B(x) = 0$. For $m_{\mathrm{b}\ell}$ and $M_{\mathrm{T2}\perp}^{221}$, this background comes primarily from events with mis-tagged jets. We estimate the shape $B(x)$ using an adaptive kernel density estimate of the shape of a control sample. This control sample is created by selecting events with one b-tagged jet and one anti-tagged jet.

We find the maximum likelihood estimator for the masses by minimizing the negative log likelihood numerically using `Minuit` [57]. To extract the final top mass estimate, we fit the distribution of masses generated from the bootstrap samples to a normal distribution. We take the fitted mean to be the measured mass, and the fitted width to be the uncertainty on the measurement.

6.3.7 Constraints

There are several different levels of constraint we can provide to the fit, depending on our goals. If our goal is to provide a precise measurement of the top-quark mass, we can provide the masses of the W boson and neutrino as inputs, fixing them to their known values. In this case, the top-quark mass is the only unknown mass to be measured. Since the $M_{\mathrm{T2}\perp}^{210}$ endpoint only depends on the W boson and neutrino masses, it does not contribute to this fit.

On the other hand, we can provide no information about any of the three masses. This allows us to explore the performance of the method in a situation

similar to a new physics mass measurement. The price for this is large statistical errors for all three measured masses.

Finally, we can do something in between. We can fix only the neutrino mass, and fit for the W boson and top-quark masses. We will present studies and results on all three variations of the fit, which we will call doubly-constrained, unconstrained, and singly-constrained, respectively.

6.4 Validation

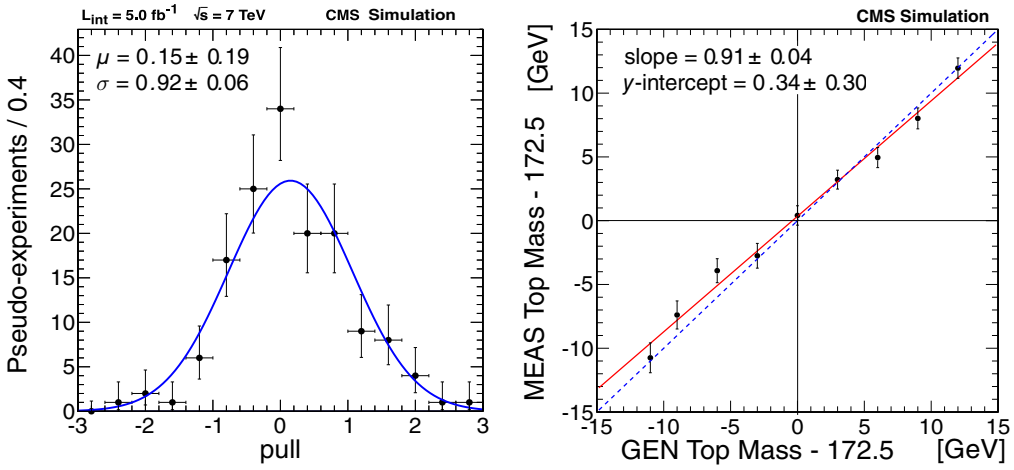
We validate the fit by using it to measure the top-quark mass in simulated events and comparing it to the top-quark mass that was used to generate the events. We do this many times, calling each simulated sample a pseudo-experiment. Each pseudo-experiment has the same number of events as are observed in the data.

By running the fit on many pseudo-experiments, we can ensure that the fit is unbiased. We frame this check for bias in terms of a pull variable, defined as

$$\text{Pull} = \frac{\text{measured} - \text{true}}{\text{uncertainty on measured}}. \tag{6.19}$$

This variable is calculated for each pseudo-experiment. If the fit is unbiased and the uncertainty estimate is accurate, the distribution of the pull variable over many pseudo-experiments should resemble a unit normal distribution. A non-zero pull mean indicates a bias in the fit, while a non-unit pull standard deviation indicates an uncertainty estimate that is too large or too small, on average.

This process is complicated by the computational expense of generating simulated events. Because of this, we must perform the validation using a limited number of simulated events. We use a sample of simulated events that is only a factor of 3 or so larger than the size of a pseudo-experiment. We generate pseudo-experiments by drawing events from this sample with replacement. This introduces



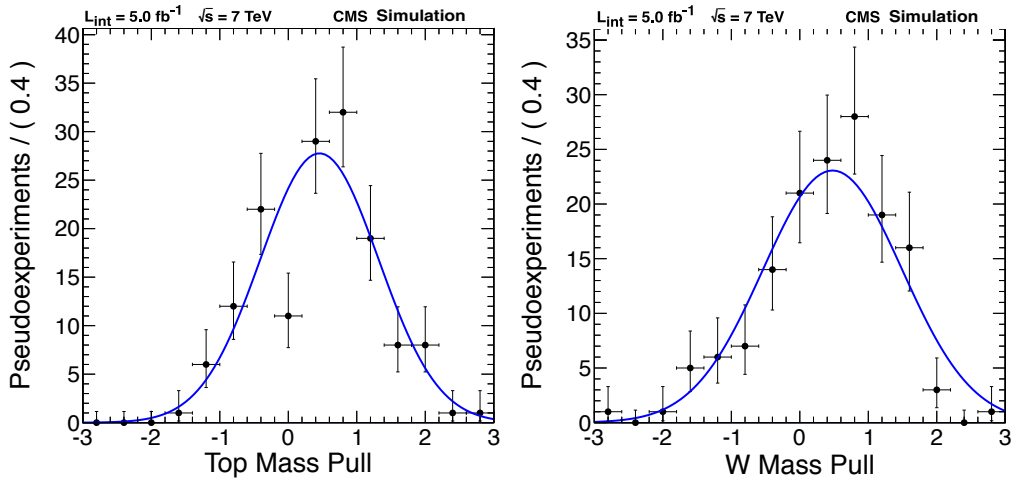
(a) (b)

Figure 6.4: (a) The pull distribution for the top-quark mass in the doubly-constrained fit over 150 simulated pseudo-experiments generated with a top-quark mass of 172.5 GeV (b) The fitted top-quark mass for simulated pseudo-experiments as a function of the top-quark mass used to generate them.

a correlation between the different pseudo-experiments, and thus the pull values from them. However, we can still correctly estimate the pull mean and standard deviation using the work of Ref. [58].

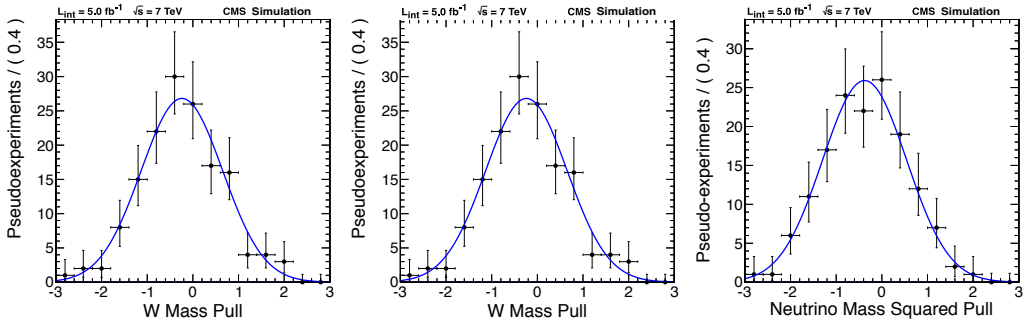
Figure 6.4(a) shows the pull distribution in the doubly-constrained fit over 150 pseudo-experiments generated with a top-quark mass of 172.5 GeV. We see that the pull distribution is approximately consistent with a unit normal, having a mean of 0.15 ± 0.19 and a standard deviation of 0.92 ± 0.06. Thus we conclude that the fit procedure is unbiased. Figure 6.4(b) shows the fitted top mass on pseudo-experiments generated with different top-quark mass values. The fact that the fitted values follow the generated values show that the procedure is unbiased regardless of the value of the top-quark mass.

Figure 6.5 and Figure 6.6 show the pull distributions for the singly-constrained and unconstrained fits, respectively. Here, we see that the biases are slightly larger, but are still small compared to the statistical uncertainty.



(a) (b)

Figure 6.5: The pull distribution for the top-quark and W boson masses in the singly-constrained fit over 150 simulated pseudo-experiments.



(a) (b) (c)

Figure 6.6: The pull distribution for the top-quark and W boson masses and neutrino mass squared in the doubly-constrained fit over 150 simulated pseudo-experiments.

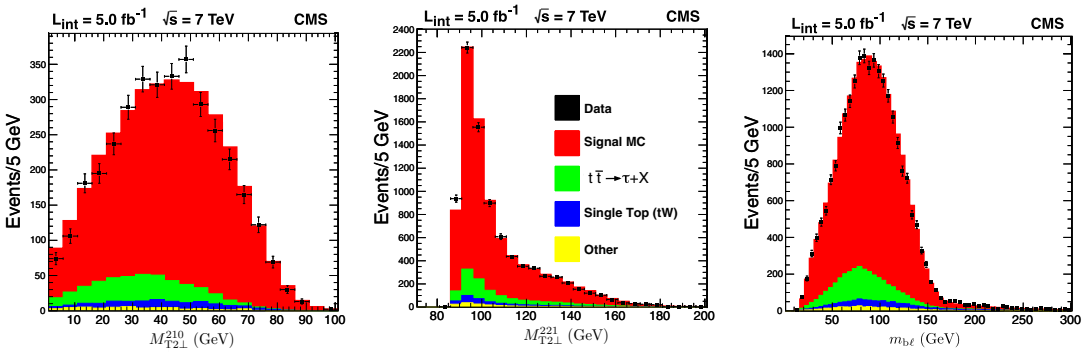


Figure 6.7: Distributions for the three kinematic variables in data and simulation.

6.5 Results

In 5.0 fb^{-1} of data taken at a center-of-mass energy of 7 TeV, we observe 8,678 events that pass the event selection. Figure 6.7 shows the $M^{210}_{\mathrm{T2}\perp}$, $M^{221}_{\mathrm{T2}\perp}$, and $m_{\mathrm{b}\ell}$ histograms in data overlayed on the histograms from simulation. Although this comparison of data to simulation does not affect the final result, it is encouraging to see good agreement.

6.5.1 Unconstrained

Figure 6.8 shows the fitted shapes to all three distributions in the unconstrained fit for one randomly-chosen bootstrap experiment. We see that the fitted shape tracks the data very well. Figure 6.9 shows the distribution of all three masses over 200 bootstrap samples. Each of these distributions is fit to a normal distribution to extract the measured mass and uncertainty. We measure $m_{\mathrm{t}} = 163 \pm 10$ GeV, $m_{\mathrm{W}} = 72 \pm 7$ GeV, and $m_{\nu}^2 = -556 \pm 473$ GeV^2. The uncertainties are statistical only; we will discuss systematic uncertainties in Section 6.6. Although these results have fairly large uncertainties, they are all roughly consistent with their world average values, which are, at the time of this writing, $m_{\mathrm{t}} = 173.5 \pm 0.6(\mathrm{stat.}) \pm 0.8(\mathrm{syst.})$ GeV, $m_{\mathrm{W}} = 80.384 \pm 0.015$ GeV, and $m_{\nu} < 0.2$ eV [59]. Note that,

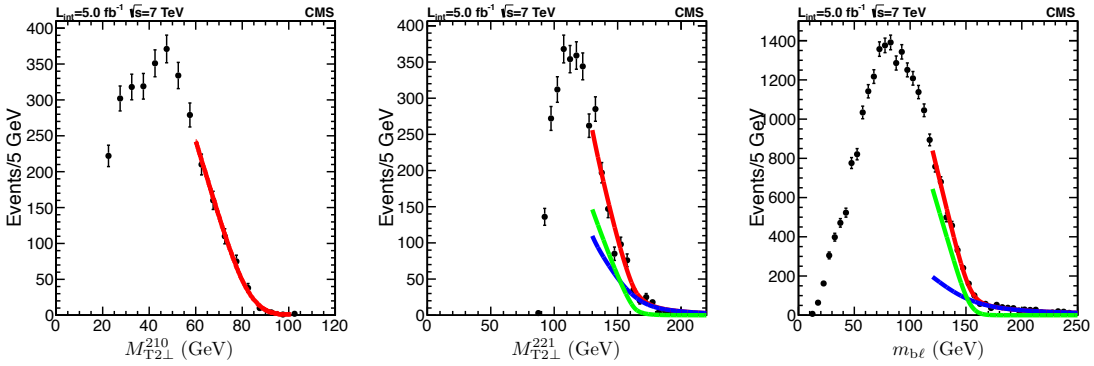
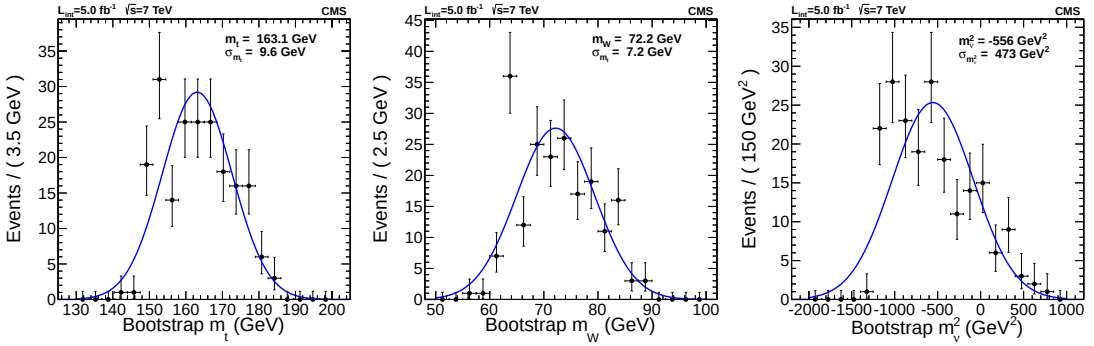


Figure 6.8: Results of the unconstrained fit on one randomly chosen bootstrap sample. The red line shows the full fit shape, while the blue and green shapes show the signal and background components, respectively.



(a) (b) (c)

Figure 6.9: The distribution of the fitted top-quark and W boson masses and neutrino mass squared over 200 bootstrap sample in the unconstrained fit.

although the fitted negative value of m_ν^2 is unphysical, the value is consistent with 0 when the uncertainty is taken into account.

6.5.2 Singly-constrained

Figure 6.10 shows the fitted shapes for the singly-constrained fit on the same bootstrap sample. In this singly-constrained fit, m_ν^2 is fixed to its known value of 0 GeV^2. Shown in Figure 6.11 are the bootstrap distributions of the W boson and top-quark masses. We measure $m_\mathrm{t} = 174.0 \pm 0.9$ GeV and $m_\mathrm{W} = 80.7 \pm 1.1$ GeV. Again, the uncertainties are statistical only. We see that constraining the neutrino

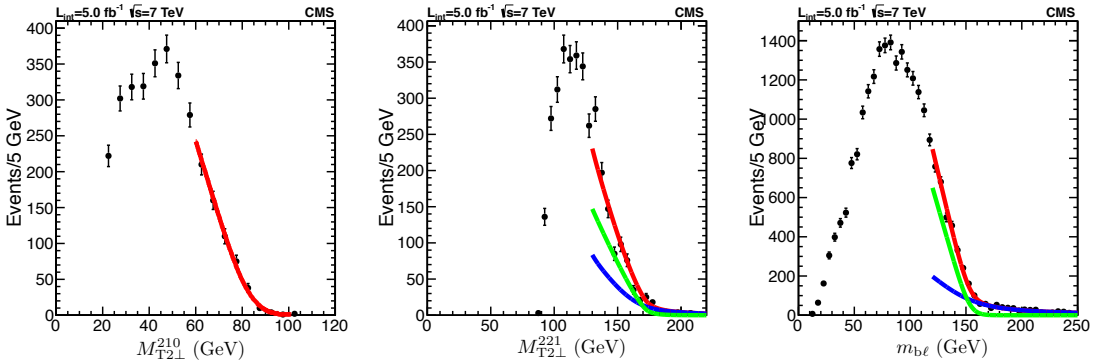
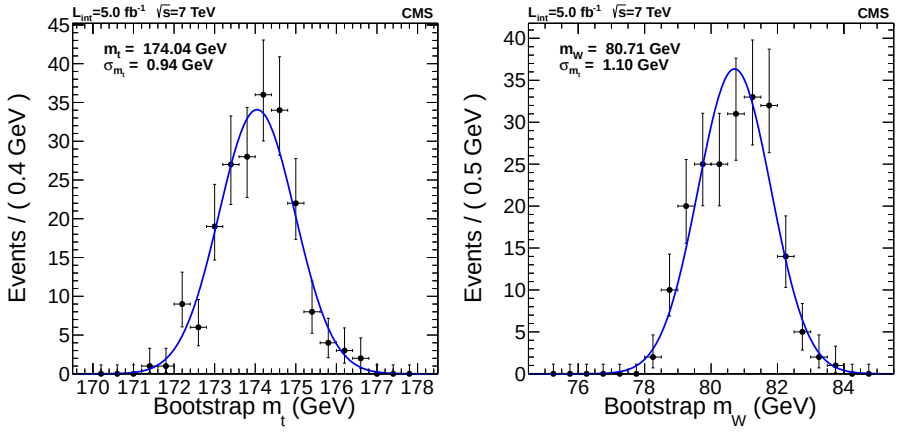


Figure 6.10: Results of the singly-constrained fit on the same bootstrap sample used in Figure 6.8. The red line shows the full fit shape, while the blue and green shapes show the signal and background components, respectively.



(a) (b)

Figure 6.11: The distribution of the fitted top-quark and W boson masses over 200 bootstrap samples in the singly-constrained fit.

mass considerably improves the precision of the measurement, and that the more precise values are still consistent with the world average values.

6.5.3 Double-constrained

Finally, the doubly-constrained result, which fixes $m_\nu = 0$ GeV and $m_{\mathrm{W}} = 80.4$ GeV, provides the most precise measurement of the top-quark mass. Since the top-quark mass does not enter into the formula for the $M_{\mathrm{T2}\perp}^{210}$ endpoint, only

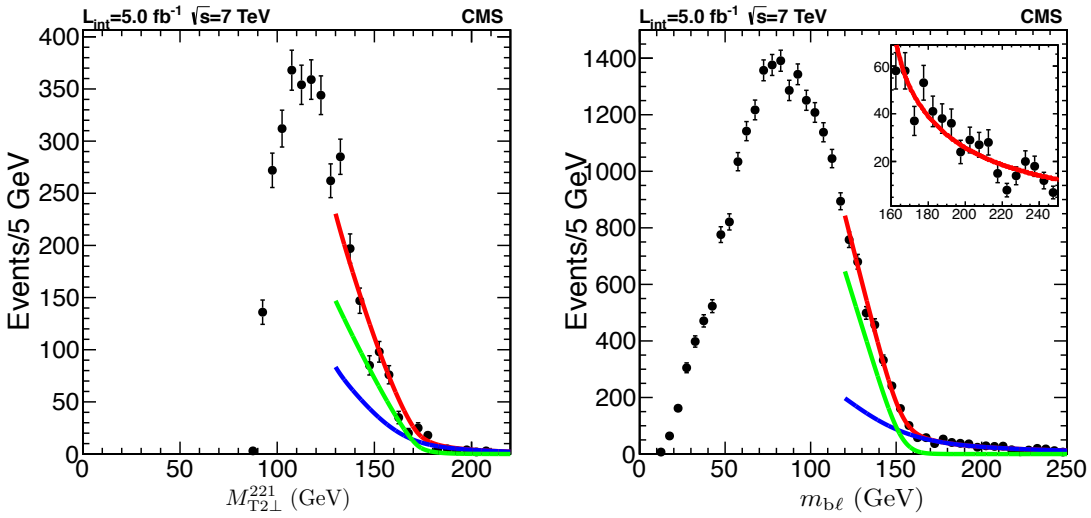


Figure 6.12: Results of the doubly-constrained fit on the same bootstrap sample used for Figures 6.8 and 6.10. The red line shows the full fit shape, while the blue and green shapes show the signal and background components, respectively.

the $m_{b\ell}$ and $M_{T2\perp}^{221}$ distributions enter into the fit.

Figure 6.12 shows the fitted shapes for theses two distributions in the same bootstrap experiment. The inset on the plot showing the $m_{b\ell}$ distribution shows the excellent agreement in the endpoint region, where the background shape modeling is very important.

Shown in Figure 6.13 is the bootstrap distribution for the top-quark mass. From this distribution, we extract our measurement, $m_t = 173.9 \pm 0.9$ GeV, where again, the uncertainty is statistical only. This result is slightly more precise than the singly-constrained result, and is in good agreement with the world average top-quark mass.

6.6 Systematic Uncertainties

We now turn to evaluating systematic uncertainties on our results. These might arise from, for example, mis-calibration of the detector, or some incorrect assump-

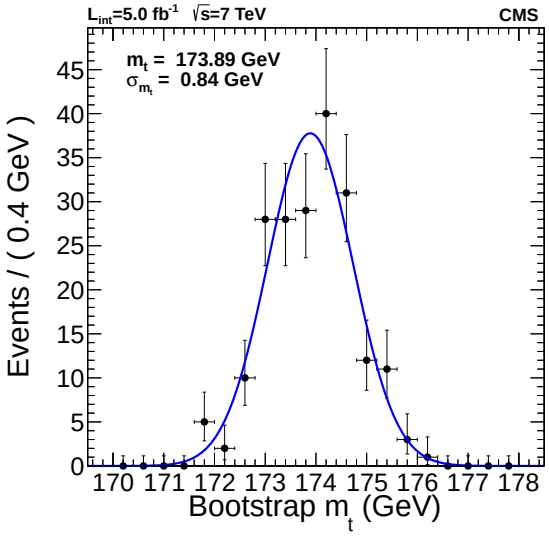


Figure 6.13: The distribution of fitted top-quark masses over 200 bootstrap samples.

tions of the analysis method. Additionally, there are a few parameters that are chosen somewhat arbitrarily. Our approach to evaluating these systematic uncertainties is to vary the parameter in question, then repeat the entire analysis with that parameter varied. This gives a good approximation of how sensitive the final result is to any particular effect.

We focus on evaluating systematic uncertainties for the doubly-constrained fit, as this top mass result is precise enough to significantly contribute to the world average. For the results with zero and one constraint, we limit ourselves to evaluating the largest sources of systematics.

6.6.1 Jet Energy Scale

One of the major sources of systematic uncertainty in any top-quark mass measurement is the jet energy scale (JES). While the relative response of the CMS detector to jets is very carefully calibrated, there is some uncertainty on the absolute calibration. Ref. [60] describes this calibration and its uncertainty. We evaluate the

effects of the JES uncertainty on the top-quark mass measurement by repeating the analysis twice: once with the JES increased by one standard deviation and once with it decreased by one standard deviation. We find deviations in the result of +1.3 GeV and −1.8 GeV, respectively, for the doubly constrained case. The deviations in the less-constrained cases can be found in Table 6.3.

Table 6.3: Systematic uncertainties due to jet energy scale uncertainty

	Unconstrained	Singly-constrained	Doubly-constrained
m_ν^2	± 200 GeV^2		
m_W	± 3 GeV	± 0.1 GeV	
m_t	± 2 GeV	$^{+1.4}_{-1.9}$ GeV	$^{+1.3}_{-1.8}$ GeV

Similarly, the resolutions for jet measurements in CMS enter into the analysis through the resolution calculation described in Subsection 6.3.3. The measurement of these resolutions in simulated data is also described in Ref. [60]. Since these measurements are performed in simulated data, there is some uncertainty when using them on real data. We account for this by varying the jet p_T resolutions up and down by 10% and re-doing the analysis for each variation. In the doubly-constrained fit, we find the uncertainty due to this effect to be ± 0.5 GeV.

6.6.2 Fit Range

In Subsection 6.3.1, the approximation of the signal shape as a straight line is described. The straight-line approximation does not apply across the entire range of the variables, as can be easily seen in Figure 6.7. Thus, a range must be chosen in which this approximation is appropriate. Deviations of the shape from a straight line introduce a dependence of our final result on the exact fit range chosen.

The nominal value of the fit range was chosen to minimize the dependence of the result on the choice of fit range. To evaluate a systematic uncertainty based on

this choice, we vary separately each end of the three fit ranges by ±10 GeV. We take the maximum deviation from the nominal result obtained under such a variation as the systematic uncertainty. Table 6.4 shows the systematic uncertainties due to the choice of fit range.

Table 6.4: Systematic uncertainties due to the choice of fit range

	Unconstrained	Singly-constrained	Doubly-constrained
m_ν^2	±590 GeV^2		
m_{W}	±8.3 GeV	±0.6 GeV	
m_{t}	±11 GeV	±0.6 GeV	±0.6 GeV

6.6.3 Background Modeling

The AKDE described in Section 6.3.2 represents a choice of how to model the mis-tag background given the control sample of tag anti-tag events. While the AKDE was found to give the least bias in validation tests, there are several other reasonable choices for modeling the background. To evaluate a systematic uncertainty due to the choice of background model, we re-do the analysis with several alternative background models.

The AKDE is the only viable non-parametric model. The alternatives consist of fitting parametric models to the control sample, then using that model as $B(x)$ when fitting. We re-performed the analysis with three different background parameterizations:

- $Ae^{-Bx} + C$
- $Ax^3e^{-Bx} + C$
- $Ax^3 + Bx^2 + Cx + D$

We take the maximum deviation from the nominal result among fits performed with these alternative background models as the systematic uncertainty. This gives

a systematic uncertainty on the top-quark mass due to background modeling of ± 0.6 GeV in the doubly-constrained fit.

6.6.4 Efficiencies

We correct for inefficiencies in both b-tagging and lepton identification, as described in Subsection 6.3.5. The efficiency values use for this correction have some associated uncertainty.. We account for the uncertainties by varying the efficiencies up and down by 1 standard deviation. We find that the efficiency uncertainty contributes at most $^{+0.1}_{-0.2}$ GeV to the top-quark mass uncertainty in the doubly-constrained fit.

6.6.5 Simulation Uncertainties

The final source of systematic uncertainty is somewhat subtle. We validate the method by applying it on simulated data, as described in Subsection 7.5.6. However, this simulation is subject to uncertainties that arise from calculations in non-perturbative quantum chromodynamics. The principle uncertainty arises from considering the amount of color reconnection of the top quarks to the proton beam remnants. We evaluate this by performing the validation on two different simulated data samples with different amounts of color reconnection, implemented as the `PYTHIA` tunes `PERUGIA2011` and `PERGUIA2011NOCR` [61]. We find that the average difference in the measured top-quark mass between the two tunes is 0.6 GeV. We take this as the systematic uncertainty.

6.7 Conclusion

Table 6.5 shows the results of this analysis with all systematic uncertainties included. The importance of these results is two-fold. First, we have shown that, with no information about any of the masses, we can measure all three masses with a precision on the order of 10%. We have also shown that, after adding our knowledge of the neutrino and W boson masses, this method gives a top-quark mass measurement with a precision that is comparable to other methods. Because the analysis uses a very different technique than the traditional ones, it may prove to have only small correlation with other measurements on the same data, and may contribute significantly to the world-average top-quark mass measurement.

Table 6.5: Final results of the analysis. The first uncertainty is statistical, while the second is systematic.

	Unconstrained	Singly-constrained	Doubly-constrained
m_ν^2	$-556 \pm 473 \pm 622$ GeV^2		
m_W	$72 \pm 7 \pm 9$ GeV	$80.7 \pm 1.1 \pm 0.6$ GeV	
m_t	$163 \pm 10 \pm 11$ GeV	$174.0 \pm 0.9^{+1.7}_{-2.1}$ GeV	$173.9 \pm 0.9^{+1.7}_{-2.1}$ GeV

CHAPTER 7

SEARCHES FOR NEW PHYSICS USING KINEMATIC ENDPOINT VARIABLES

We have shown that kinematic endpoint variables can be used for mass measurements in a realistic setting. What else can these variables be used for? One use that comes to mind is in searches for new physics. We know from the previous chapter that we can select final states in which SM processes have a definite upper endpoint. This creates a region above the endpoint where we, in a perfect experiment, expect no SM events. Therefore, if we observe any events in that region, they may come from some form of new physics. Of course, as we have already seen, our experiment is not perfect, and SM events can yield values above the endpoint due to mis-measurements and backgrounds. However, it is possible to understand and account for SM events above the endpoint.

We saw in the top-quark mass measurement that the $M_{\mathrm{T2}\perp}^{210}$ variable has a particularly low number of events above the endpoint. We were able to make the approximation that background contributions were negligible, and the resolution of $M_{\mathrm{T2}\perp}^{210}$ was small. These factors make this variable is a natural choice to use in search for new physics.

In this chapter, we present an analysis that takes advantage of this idea. We search for new physics in events with topologies similar to the 210 subsystem: an opposite-sign dilepton pair and missing transverse momentum. After some preselection requirements, the primary backgrounds come from $\mathrm{t\bar{t}} \to \mathrm{bW^+\bar{b}W^-} \to \mathrm{b}\ell^+\bar{\nu}\bar{\mathrm{b}}\ell^-\nu$ and $\mathrm{WW} \to \ell\nu\ell\nu$ events, both of which exhibit the same well-defined upper endpoint in the $M_{\mathrm{CT}\perp}$ variable at m_{W}. By looking for events above this endpoint, we can perform a very generic search for new physics.

This search is sensitive to any new physics that produces events containing opposite-sign dilepton pairs and $\not{p}_{\mathrm{T}}$ that do not respect the kinematic endpoint. This can come about either because of the presence of particles with different masses than those in the SM or because the events do not match the event topology that the variable assumes. This type of search has been performed before, albeit with different background estimation techniques, in generic hadronic searches [40], searches for scalar bottom quarks [41], and searches for sleptons and charginos [42]. The idea has appeared in the theory literature a number of times [39, 62–64].

Although the search is generic, we focus on supersymmetric models that have been less-explored at the LHC: those with weakly-produced SUSY particles. In particular, the results of this search are interpreted in terms of two simplified models [65] describing chargino and slepton pair production. The first model, shown in Figure 7.1(a), describes electroweak chargino pair production. We assume that sleptons and sneutrinos are light, so that a chargino can decay via either a slepton or a sneutrino. This sparticle then decays again to give a final state with a neutrino, a charged lepton, and an LSP. While Figure 7.1(a) shows one chargino decaying to a slepton and the other to a sneutrino, the decays of the two charginos are in fact independent of each other. Each chargino decays to a slepton or a sneutrino with equal probability. Thus, we have several different decay modes of the pair of charginos, but all give the same final state: two charged leptons, two neutrinos, and two LSPs. For simplicity, the slepton and sneutrino masses are fixed to be equal to the average of the chargino and the LSP masses, while limits are set as a function of the chargino and LSP masses.

In this model, we make the assumption of flavor democracy, so that a chargino is equally likely to decay to any flavor of slepton or sneutrino. Taking into account that this analysis will only observe τs when they decay leptonically (again flavor-

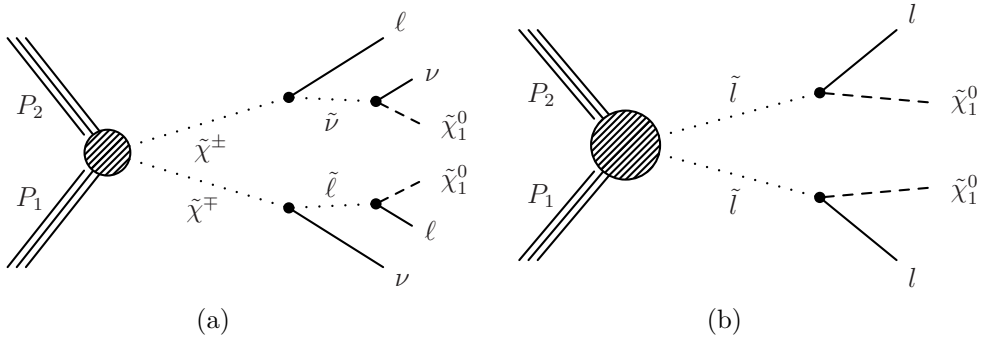


Figure 7.1: Diagrams showing the simplified models used in this analysis. (a) A model describing chargino pair-production and decay via sleptons and sneutrinos. This model yields a final state with two leptons and $\not{p}_{\mathrm{T}}$. (b) A model describing light slepton pair-production with each slepton decaying to a lepton and an LSP. This model also gives a final state with two leptons and $\not{p}_{\mathrm{T}}$.

democratically), we expect this model to produce equal numbers of events with a same-flavor lepton pair (ee or $\mu\mu$) as with an opposite-flavor pair (eμ).

The second model, shown in Figure 7.1(b), describes the electroweak production of light sleptons (smuons or selectrons). Each slepton decays to its partner lepton and an LSP. In this model, the produced sleptons must be the same flavor in order to avoid lepton flavor violation. Thus, the lepton pair in the final state is always same-flavor. The selectron and smuon masses are set to be equal to each other, and we vary both the slepton and the LSP mass in the interpretation.

7.1 Event Selection

We require events to have exactly two opposite-sign leptons which pass all selection criteria. Selected leptons are required to be isolated and have $p_{\mathrm{T}} > 20$ GeV. Here, an isolated lepton is defined as having pileup-corrected relative isolation in a cone of $\Delta R \leq 0.3$ less than 0.15.

We also require $\not{p}_{\mathrm{T}} > 60$ GeV. The main purpose of the $\not{p}_{\mathrm{T}}$ cut is to discriminate against Drell-Yan production with fake MET due to mis-measured ISR jets. To further suppress on-shell Z boson production, we veto same-flavor events where

the dilepton mass lies within 15 GeV of the Z mass (91 GeV). To reduce $t\bar{t}$ backgrounds, we veto events that contains b jets. B jets are identified using the Combined Secondary Vertex b-tagger at the medium working point, and must have $p_T > 30$ GeV and $|\eta| < 2.5$.

Table 7.1: Data and simulation event yields after preselection cuts. Events are split depending on the leptons in the final state. Events with two electrons or muons are considered to be same flavor, while events with one muon and one electron are considered to be opposite-flavor.

Background Type	Opposite-Flavor	Same-Flavor
Top	7787	5840
WW	3061	2460
WZ	169	204
ZZ	3	67
Z/γ^*	1354	4681
W	2014	340
Total Simulation	13,398	13,592
Data	13,152	15,204

Table 7.1 shows event yields in data and simulation after applying the preselection. As can be seen, major backgrounds to this search are those which include two isolated leptons plus real MET. These include top-quark pair production with a dileptonic decay. Also important are diboson production processes with decays to two leptons and neuttrinos. The most prominent of these is WW $\rightarrow \ell\nu\ell\nu$. Since all these processes have two isolated leptons and real $\not{p}_T$, the preselection does nothing to discriminate against them.

Less important are those backgrounds which are effectively discriminated against by the preselection. These include Z/γ^* production, which is reduced by both the MET requirement and the Z-mass veto. Backgrounds with only one lepton can contribute when they have real $\not{p}_T$ and either a jet faking a lepton or a non-isolated lepton from a jet. The two primary processes that contribute to this are W boson production, because of its large cross-section, and semileptonic $t\bar{t}$ decays, because

the b jets produce non-isolated leptons more frequently than light jets. Contributions from these processes are controlled by lepton quality and isolation cuts.

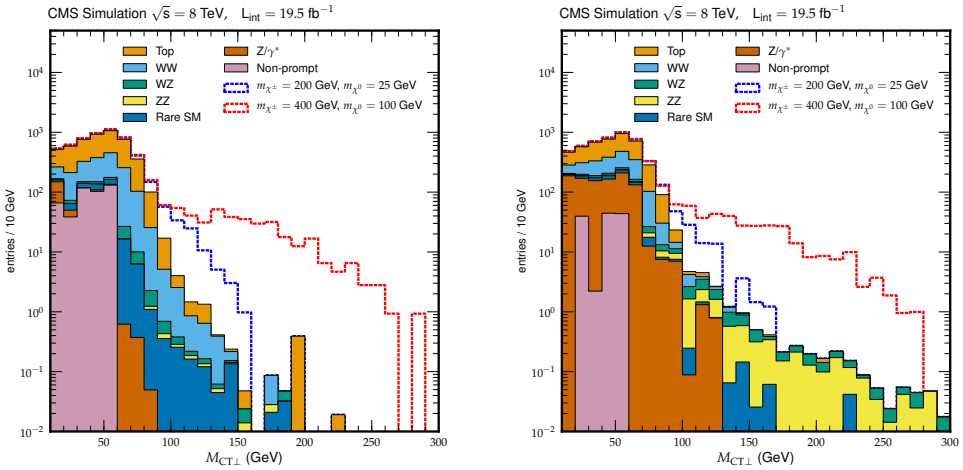
7.2 Analysis Strategy

In this analysis, we use $M_{\mathrm{CT}\perp}$ to discriminate against backgrounds containing two W bosons that both decay leptonically. Taking the two leptons as our visible particles, Equation 5.19 gives

$$M_{\mathrm{CT}\perp}^{\max} = \frac{m^2(W) - m^2(\nu)}{m(W)} = m(W). \tag{7.1}$$

We can search for new physics by looking for events beyond the $M_{\mathrm{CT}\perp}$ endpoint. Figure 7.2 shows the $M_{\mathrm{CT}\perp}$ distribution from Monte Carlo simulation. We can see that although the backgrounds do extend past the expected endpoint, they fall off quickly. New physics signals will generally involve different particle masses, and may extend to higher $M_{\mathrm{CT}\perp}$ values.

We split the analysis into events with opposite-flavor and same-flavor lepton pairs, as the backgrounds and signal models populate the two channels differently. The analysis focuses on the *shape* of the $M_{\mathrm{CT}\perp}$ distribution in these channels, which is expected to be different for a potential signal than for the background. This is framed in terms of a likelihood describing the shape of the $M_{\mathrm{CT}\perp}$ distribution. We first derive templates describing the shape of the $M_{\mathrm{CT}\perp}$ distribution for each background. We then construct a model probability density function (PDF) for the observed $M_{\mathrm{CT}\perp}$ values using a sum of the background templates. The normalizations of the templates are unknown, and are left as free parameters in the model. We then ask the question: "Can the data be described well by this background-only model?" To quantify this, we perform a hypothesis test, where the null hypothesis is that the data is described by the sum-of-backgrounds model.



(a) (b)

Figure 7.2: The $M_{\mathrm{CT}\perp}$ distribution from Monte Carlo simulation in the (a) opposite and (b) same-flavor channels. Filled histograms show SM backgrounds, while the dashed histograms show two different mass points in the simplified model shown in Figure 7.1(a).

We perform the hypothesis test using the Anderson-Darling goodness-of-fit test statistic [66], which has the virtue of requiring no specific alternative hypothesis. This test is described in more detail in Subsection 7.5.3.

When setting limits on specific signal models, we have a specific alternative hypothesis to test. The alternative hypothesis supposes that the $M_{\mathrm{CT}\perp}$ PDF is given by the sum of the background templates, *plus* a signal template, which is derived from simulation of a specific signal model. In this case, we can use the more-powerful profile likelihood ratio test statistic. We then set limits using the standard $\mathrm{CL_S}$ procedure [67].

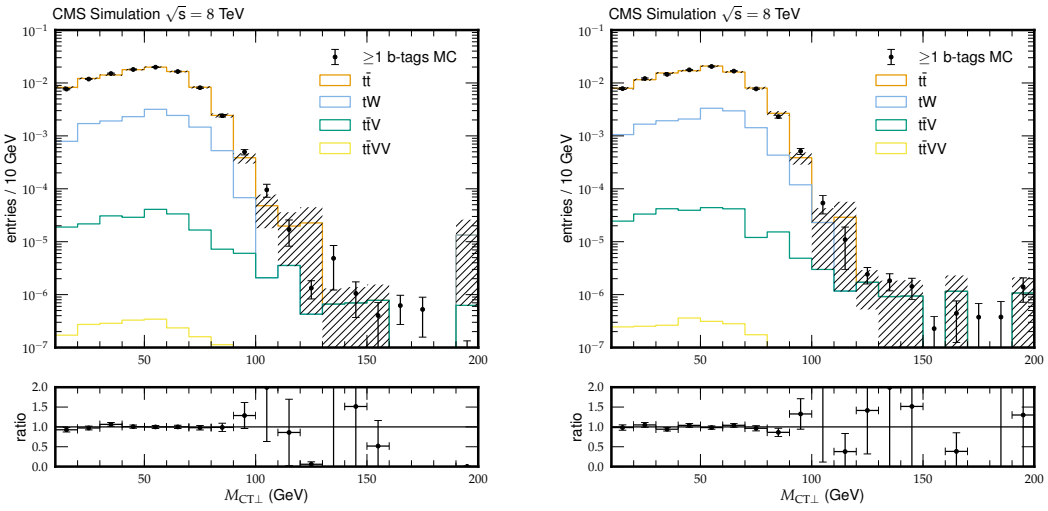
With this analysis strategy, we do not need to know the normalization of the backgrounds, but the analysis hinges critically on deriving background templates which accurately describe the *shapes* of the backgrounds. Thus, the major focus of the analysis is on deriving the templates, validating them, and assigning systematic

uncertainties on the shapes.

7.3 Background Modeling

We derive the shapes of the background $M_{\mathrm{CT}\perp}$ distributions from either control regions in data or from simulation. Where possible, we use data control regions in which one type of background dominates. To validate these, we look at both the control and signal regions in simulation to ensure that the control region describes the signal region well.

In order for a control region to correctly describe the $M_{\mathrm{CT}\perp}$ distribution of a background process, the process of selecting events for the control region must not distort the shape of the $M_{\mathrm{CT}\perp}$ distribution. As a counter-example, consider the logical choice of a control region for processes containing Z bosons. One might choose the region where the dilepton invariant mass is consistent with a Z boson decay (inverting the Z veto in the event selection). However, the dilepton invariant mass is correlated with $M_{\mathrm{CT}\perp}$, so this control region would not have the same $M_{\mathrm{CT}\perp}$ distribution as the signal region. In this case and several others, we use simulation to predict the background shape by looking at simulated background events in the signal region. We can verify that the simulation describes the data well by comparing data and simulation in a control region. In this case, the control region does not need to accurately reflect the signal region $M_{\mathrm{CT}\perp}$ distribution, as we are only trying to validate the simulation. The control region simply needs to be orthogonal to the signal region.



(a) (b)

Figure 7.3: Comparison of the $M_{\mathrm{CT}\perp}$ shapes of the top control region versus the true top shapes in Monte Carlo simulation in events where the two leptons are of the opposite (a) and same (b) flavor. The simulated signal region histograms from different processes are stacked and their sum normalized to one.

7.3.1 Top

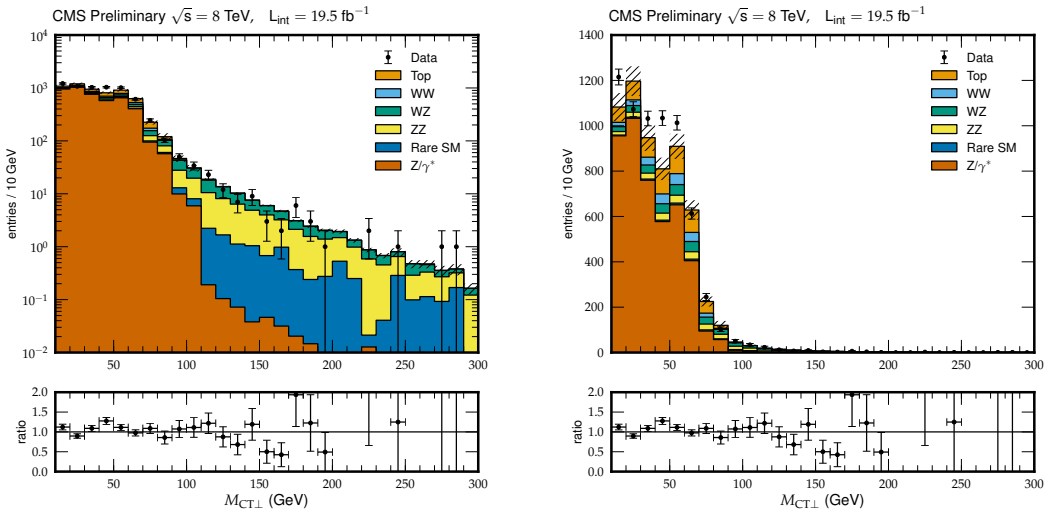
Our backgrounds from top events include $t\bar{t} \to bW^+\bar{b}W^- \to b\ell^+\overline{\nu}\bar{b}\ell^-\nu$, $tW \to bWW \to b\ell\nu\ell\nu$, and rare processes which produce $t\bar{t}$ in association with one or more vector bosons. We predict the shape of these combined backgrounds from a control sample. To obtain the control sample, we take advantage of the b-jets in these processes by inverting our b-jet veto. Thus, the control sample is composed of events that have at least one b-tagged jet. We use the same tagger that is used in the signal selection. Figure 7.3 shows the shape of this control sample overlaying the $M_{\mathrm{CT}\perp}$ distribution in the signal region for the sum of all top-like events in Monte Carlo simulation. We see that the agreement between the two shapes is very good. The inset shows the ratio of bin counts in the signal region truth to those in the control region.

7.3.2 Diboson and Rare SM Processes

Backgrounds arise from the WW process via the dileptonic decay WW $\rightarrow \ell\nu\ell\nu$, which looks very similar to signal. They can also arise from WZ production via the decay WZ $\rightarrow 3\ell\nu$ if one lepton is undetected or fails the quality cuts. The third diboson process, ZZ, also contributes in the same flavor channel. This occurs when an off-shell Z boson decays leptonically and the other decays via neutrinos, creating an off-Z dilepton pair plus $\not{p}_{\mathrm{T}}$. These diboson backgrounds offer no data-driven control sample, as they are only distinguished from signal by their $M_{\mathrm{CT}\perp}$ distributions.

We get the template for the combined WW, WZ, and ZZ shape from Monte Carlo simulation. We also add rare processes including three vector bosons and Higgs boson decays via WW to this shape. The relative amount of each process in the template is fixed, but the overall template normalization is allowed to vary. This is equivalent to fixing the ratios of the WW, WZ, ZZ, VVV, and H $\rightarrow$ WW cross sections. We account for uncertainties in the relative cross-sections by adding shape systematics created by shifting the cross-section for each component up and down by 10% for the diboson processes and by 50% for the VVV and H $\rightarrow$ WW processes.

We perform validation of the Monte Carlo simulation to confirm that the simulation describes the data well. We do this by looking at control samples which are orthogonal to the signal region. First, we look in on-Z events, considering only the high $M_{\mathrm{CT}\perp}$ region where ZZ $\rightarrow \ell\ell\nu\nu$ dominates, shown in Figure 7.4. We see that the agreement is very good in both shape and normalization. Second, we can isolate a sample of mainly WZ events by requiring the existence of third lepton and at least one opposite-sign same-flavor lepton pair with an invariant mass near the Z boson mass. The $M_{\mathrm{CT}\perp}$ distribution for this sample is shown in Figure 7.5.



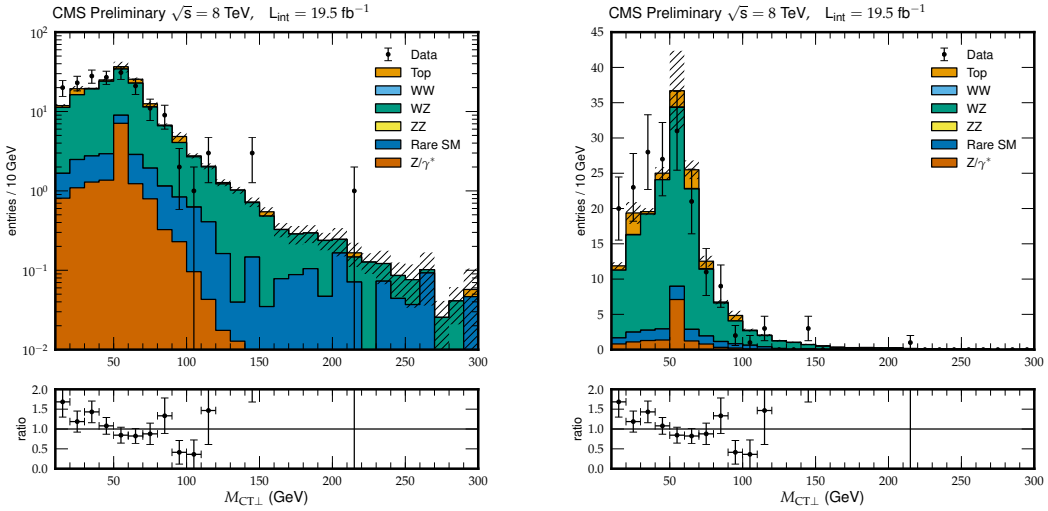
(a) (b)

Figure 7.4: Comparison of data and Monte Carlo simulation with all preselection cuts applied, but the Z mass veto inverted in the same flavor channel. In the low $M_{\mathrm{CT}\perp}$ region, the Z/γ^* background dominates. Discrepancies from this plot are used to assign a shape systematic to the Z template. In the high $M_{\mathrm{CT}\perp}$ region, the ZZ background dominates, and we see that the agreement is quite good. Y-axis scales are a) log and b) linear.

We again see good agreement.

Validation of the WW simulation is somewhat more complicated. We take advantage of the fact that W and Z bosons have very similar leptonic decays, except a W boson decay produces a neutrino instead of the Z boson's second lepton. Therefore, if we take a leptonic Z decay and treat one of the leptons as a neutrino, we have created a decay that looks very much like a leptonic W decay.

To validate the WW Monte Carlo simulation, we select a relatively pure sample of WZ events by requiring events to have three leptons and $\not{p}_{\mathrm{T}} > 30$ GeV. To ensure that we do not calculate $M_{\mathrm{CT}\perp}$ from two leptons that both come from a Z boson, we require that an opposite-sign, opposite-flavor pair exists. This is the pair that we use in the $M_{\mathrm{CT}\perp}$ calculation. The third lepton is removed from the event,



(a) (b)

Figure 7.5: Comparison of data and Monte Carlo simulation with all preselection cuts applied. We require three leptons, two of which form an invariant mass consistent with a Z boson. In this region WZ background is dominant, and we see good agreement between data and simulation. Y-axis scales are a) log and b) linear.

and its four-momentum is added to the $\not\!\!p_{\mathrm{T}}$ vector in order to simulate a neutrino. Having done this, we re-calculate $M_{\mathrm{CT}\perp}$ and $\not\!\!p_{\mathrm{T}}$. We must then re-scale the $M_{\mathrm{CT}\perp}$ values to account for the different endpoint that arises when M_{W} is replaced with M_{Z} in Equation 7.1. To do this, we simply multiply all $M_{\mathrm{CT}\perp}$ values by $M_{\mathrm{W}}/M_{\mathrm{Z}}$, since we can derive from Equation 5.43 that the entire $M_{\mathrm{CT}\perp}$ distribution scales linearly with the parent mass. Finally, we apply our signal region cuts, and are left with events that resemble WW events. We denote this the $3-1$ lepton control region.

Figure 7.6 shows a comparison of the $M_{\mathrm{CT}\perp}$ shapes of the $3-1$ lepton control region and the WW template in Monte Carlo simulation. We calculate a re-weighting factor for each bin needed to make the two distributions in the figure match. These factors are applied to the $3-1$ lepton control region in data. Figure 7.7(a) shows the comparison between the re-weighted control region in data and the WW Monte Carlo shapes, while Figure 7.7(b) shows a shape comparison between the unweighted data control region and the simulated control region. We see reasonable agreement, but the control region suffers from a low number of events. For this reason, we calculate a bin-by-bin shape systematic on the WW template from Figure 7.7(a). For each bin, we take the larger of the discrepancy between the re-weighted control region and the WW template and the statistical uncertainty on the control region.

7.3.3 $\mathbf{Z/\gamma^*}$

We take the shape for Z and Drell-Yan backgrounds from Monte Carlo simulation. These events contain no real $\not\!\!p_{\mathrm{T}}$, and the simulation does not model fake $\not\!\!p_{\mathrm{T}}$ well, as can be seen in Figure 7.8. This figure shows the $\not\!\!p_{\mathrm{T}}$ distribution in events passing signal selections other than the $\not\!\!p_{\mathrm{T}}$ and invariant mass requirements. Then, instead

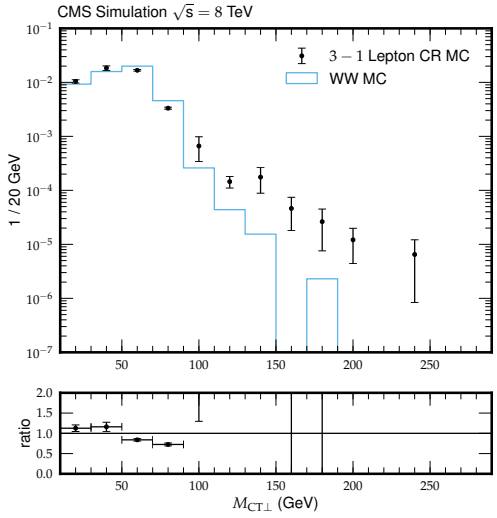
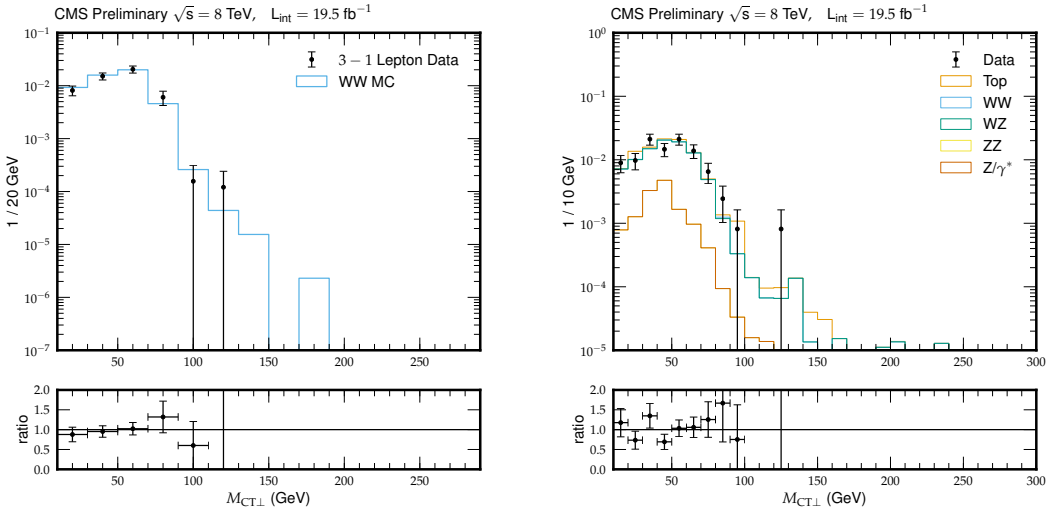


Figure 7.6: Comparison of the $3-1$ lepton control region $M_{CT\perp}$ shape in Monte Carlo to that of WW Monte Carlo. Reweighting factors calculated from this comparison are applied to the $3-1$ lepton control region in data.



(a) (b)

Figure 7.7: a) Comparison of the reweighted $3-1$ lepton control region $M_{CT\perp}$ shape in data to that of WW Monte Carlo. Discrepancies in this comparison are assigned as shape systematics on the diboson template. b) Shape comparison of the unweighted $3-1$ lepton control region in data and the $3-1$ lepton control region in Monte Carlo.

of rejecting events where the dilepton invariant mass is near the Z mass, we require the invariant mass to be within 15 GeV of the Z mass. This sample of events is dominated by processes with Z bosons.

To address the discrepancy in the $\not{p}_{\mathrm{T}}$ distribution, we perform a binned reweighting of the Z/γ^* simulation such that the $\not{p}_{\mathrm{T}}$ distribution in this control region agrees with data. We first subtract the small Monte Carlo expectations for non-Z backgrounds from the data histogram in this control region. We then determine the binning. We start with 10 GeV-wide bins and iteratively remove bin boundaries until all bins have at least 5 events in both data and simulation. We calculate a correction factor for each bin such that simulation = data in each bin. The entire Z/γ^* simulation sample is then reweighted by the factor corresponding to the $\not{p}_{\mathrm{T}}$ bin each event falls into. All plots and values in this chapter other than Figure 7.8 have this correction applied.

To validate this corrected simulation, we look again at Figure 7.4, which shows the data and simulation $M_{\mathrm{CT}\perp}$ distributions in the on-Z control region after the corrections. We now consider only the low-$M_{\mathrm{CT}\perp}$ region where the Z background dominates, rather than the higher-$M_{\mathrm{CT}\perp}$ region where diboson events are more prominent. We see that there is some disagreement between the data and simulation, even after the correction.

Given that the Monte Carlo simulation does not perfectly match the data in the control region, we apply systematic uncertainties to the signal region template that we obtain from this simulation. To do this, we assign the average relative discrepancy between the data and simulation shapes in the control region as a systematic uncertainty in each bin of the Z/γ^* template. This gives an uncertainty of 24% on each bin.

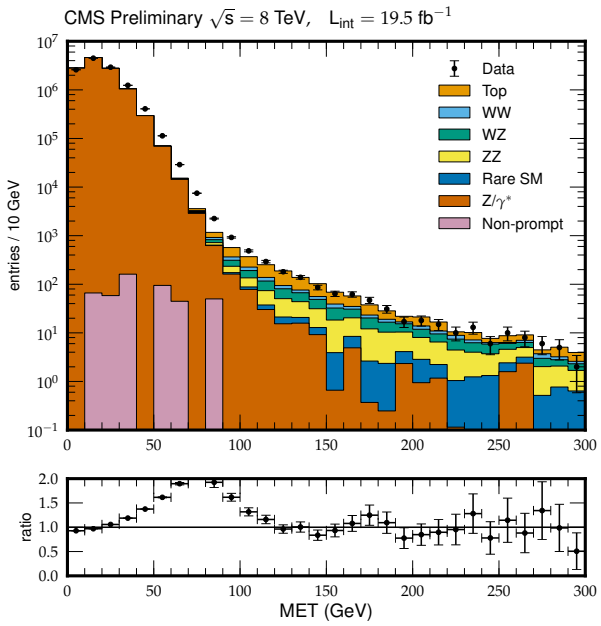


Figure 7.8: Data-Monte Carlo comparison of the $\not\!\!p_{\mathrm{T}}$ distribution in the on-Z control region. The Z Monte Carlo simulation is henceforth reweighted to correct for the discrepancy in the distributions.

7.3.4 Non-prompt Leptons

We obtain a sample of events with non-prompt leptons originating from W boson and $t\bar{t}$ decays with only one prompt lepton (semileptonic decays) by changing the relative isolation cut on one of the leptons and requiring that the two selected leptons have the same charge. The relative isolation is described in Section 7.1; we require one of the leptons to have a relative isolation between 0.2 for 0.3, while the other must have a relative isolation < 0.15 as in the signal selection. Figure 7.9 shows the shape of the $M_{\mathrm{CT}\perp}$ distribution in this control region in simulation overlaid with the simulated signal region shape. As the number of events in the simulated event sample is small, we combine the same-flavor and opposite-flavor channels into the same plot. Agreement looks good, but it is hard to draw any firm conclusions because of the small number of simulated events in the W boson sample. Thus, we conservatively assign a 30% bin-by-bin systematic uncertainty to the data-driven non-prompt template.

7.3.5 Flavor-Symmetric

In the case of slepton pair production, we expect the final state to produce only same-flavor lepton pairs. This allows us to use an alternative background modeling method for backgrounds which produce equal numbers of same-flavor and opposite-flavor events. These backgrounds are comprised of top, WW, WZ, H $\rightarrow$ WW, and VVV events. Rather than deriving separate templates for these backgrounds as previously described, we can combine all of these backgrounds into a single template, which we model using a control sample of events with opposite-flavor lepton pairs. By combining several backgrounds into a single template, we reduce the number of unknown normalizations in the statistical model, increasing the power of the analysis. Thus, when considering signals that produce only same-

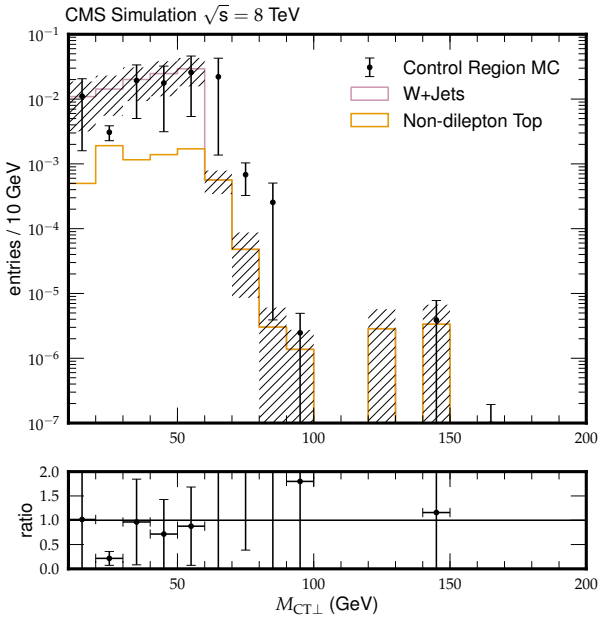


Figure 7.9: Comparison of the $M_{\mathrm{CT}\perp}$ shapes of the non-prompt control region versus the true W +Jets and semileptonic $\mathrm{t\bar{t}}$ shape in Monte Carlo simulation. Both the signal and control region distributions are normalized to 1.

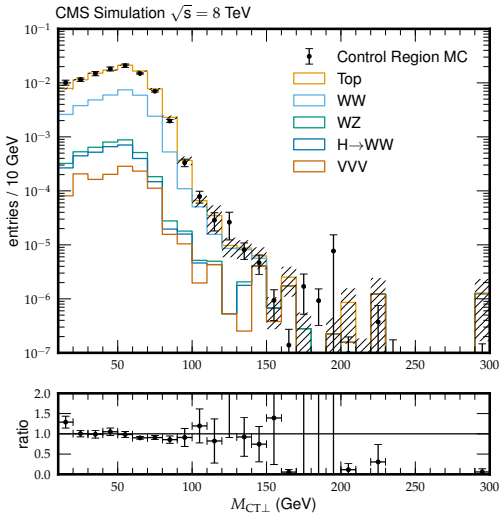


Figure 7.10: Simulated closure test of the flavor-symmetric background template. Points show the histogram derived from the opposite-flavor control region, while the line histograms show the simulated same-flavor signal region for flavor-symmetric processes.

flavor events, it is preferable to use the combined template for flavor-symmetric backgrounds. A result using this template also provides a cross-check for the other background modeling methods, at least in the same-flavor channel.

Figure 7.10 shows a closure test in simulation which overlays the shape of the opposite-flavor control sample with the same-flavor signal region simulation of the flavor-symmetric backgrounds. We see that agreement is generally good.

When using this method to model flavor-symmetric backgrounds, we must modify slightly the diboson and rare SM template described in Subsection 7.3.2. Now, this template must model only the subset of these processes that is *not* flavor-symmetric. These are WZ events where the selected leptons both come from the Z boson and ZZ events.

7.4 Signal Modeling

We will interpret the results of this search in terms of limits on the two signal scenarios shown in Figures 7.1(a) and 7.1(b). We model these processes by Monte Carlo simulation using `PYTHIA` [49]. Unlike the background processes, the signal models have their absolute normalization fixed to next-to-leading order cross-sections calculated using `Prospino 2.1` [68]. In these samples, the CMS detector is simulated using the CMS fast simulation program [69] instead of `GEANT4`.

Because the normalization is fixed and there is no way to cross-check the simulation against the actual processes in data, several systematic uncertainties are applied to these templates. First, the normalization is affected by uncertainties in integrated luminosity (4.4%), trigger efficiencies (5%), lepton identification and isolation efficiencies (2%), and b-jet veto efficiency (5%). We also take into account the effect of jet energy scale uncertainty on the shape of the $M_{\mathrm{CT}\perp}$ distribution.

7.5 Statistical Procedure

In order to apply statistical procedures to the analysis, we must first define the likelihood of the data under the various hypotheses we wish to test. Once that is done, there are several techniques available to interpret the data. We test the consistency of the data with the background-only hypothesis using an Anderson-Darling test, and set limits on specific signal models using the profile-likelihood ratio test statistic and the $\mathrm{CL_S}$ procedure.

7.5.1 Statistical Model

We model the $M_{\mathrm{CT}\perp}$ distribution of the observed data as a sum of the background distributions. Since we do not have parametric shapes for the background distri-

butions, we use a binned approach. We group the data into 10 GeV-wide bins in $M_{\mathrm{CT}\perp}$. We also split the analysis into the opposite-flavor and same-flavor channels. Here, we describe the procedure for a single channel. Since the channels are independent of each other, combining two channels is straightforward.

Let us label bins with the index i. The likelihood of observing y_i events in bin i is given by the Poisson distribution:

$$P(y_i) = \mathrm{Poisson}(\lambda_i), \tag{7.2}$$

where λ_i is the expected yield in this bin, and is in general a function of unknown parameters. The likelihood of the entire data is given by a product over all bins

$$\prod_i P(y_i). \tag{7.3}$$

The expected yield in a bin is just the sum of the expected yields of the processes that contribute events in that bin. Letting the index a denote a particular process,

$$\lambda_i = \sum_a N_a b_{ai}. \tag{7.4}$$

N_a is the normalization for background a, while b_{ai} is the yield in bin i for that background, normalized such that

$$\text{bin width} \times \sum_i b_{ai} = 1. \tag{7.5}$$

In many cases, N_a is a free parameter. However, we will also encounter cases where N_a is expressed in terms of some other parameter and has some associated uncertainty. In that case, the value of N_a is considered to be an observation from a normal distribution with unknown mean η_a so that

$$N_a \sim \mathrm{Normal}(\eta_a, \sigma_{N_a}) \tag{7.6}$$

and

$$\lambda_i = \sum_a \eta_a b_{ai}. \tag{7.7}$$

Here and throughout, we use the symbol $\sim$ to indicate "is distributed as".

If we ignore systematic uncertainties, b_{ai} is taken directly from the template for background a. However, there are several systematic uncertainties that we would like to include. Therefore, we parameterize b_{ai} as

$$b_{ai} = \alpha_i \beta_{ai} \gamma_{ai} \sigma_{ai}. \tag{7.8}$$

Here, σ_{ai} is the yield from the template, while α, β, and γ are random variables with expectation value 1, each of which accounts for a different type of systematic uncertainty.

The term α_i describes the statistical uncertainty on all of the b_{ai} in bin i. This accounts for the fact that the templates come from either data control samples or finite-size simulated event samples. We follow the approach of Ref. [70]. We can constrain the true mean yield, τ_{ai}, from our raw observation of the yield in one sample, which we will call m_{ai}. This observation has likelihood

$$m_{ai} \sim \text{Poisson}(\tau_{ai}) = \text{Poisson}(\alpha_{ai} m_{ai}), \tag{7.9}$$

where we have parameterized τ_{ai} in terms of m_{ai}.

In principle, this gives us a term α_{ai} for each bin and each process. However, this quickly becomes computationally intractible as we add processes. Instead, we note that the sum of variables that are Poisson distributed is also Poisson distributed. We then have

$$m_i = \sum_a m_{ai} \sim \text{Poisson}(\alpha_i m_i). \tag{7.10}$$

Framing the constraint in this way gives us only one unknown parameter per bin, while still preserving information about the statistical uncertainty on the templates. Note that m_i is the sum of *unweighted* event counts in all templates. For weighted templates, this corresponds to the sum of weights squared.

The next term, β_{ai}, is fairly straightforward. It describes bin-by-bin uncertainties on b_{ai}, assuming that each bin is independent from the others. We call this a *shape systematic*. This can be used to describe a lack of knowledge about the true shape of the $M_{\mathrm{CT}\perp}$ distribution, for example, if we suspect the simulation used to generate the template does not accurately reflect data. The uncertainty modeled by β_{ai} is given as a fractional uncertainty, so that

$$\beta_{ai} \sim \mathrm{Normal}(1, S_{ai}). \tag{7.11}$$

Where S_{ai} is the fractional uncertainty assigned to bin i. We truncate the normal distribution so that $\beta_{ai} > 0$.

The final parameter, γ_{ai}, describes correlated uncertainties on the $M_{\mathrm{CT}\perp}$ shape due to an uncertainty in an underlying parameter ζ. We call this a *histogram systematic*. This may arise if some parameter that we know affects the shape of the distribution is not fully determined, for example, the jet or lepton energy scale.

Because of the generally complicated interaction between the parameter and the shape, we cannot in general reproduce the full dependence on the parameter. Instead, we vary ζ up and down by one standard deviation and recalculate the template for each variation. We label the templates produced under these variations as σ_{ai}^{+} and σ_{ai}^{-}.

Next, let us reparameterize ζ. We define

$$\xi = \frac{\zeta - \zeta_0}{\sigma_\zeta}, \tag{7.12}$$

where ζ_0 is the nominal value of ζ and σ_ζ is its standard deviation. Thus, we can work with a variable that is normally distributed:

$$\xi \sim \mathrm{Normal}(0, 1). \tag{7.13}$$

We have a few constraints that must be satisfied. We know that the $\gamma_{ai}\sigma_{ai} = \sigma_{ai}^{+}$ when $\xi = 1$, $\gamma_{ai}\sigma_{ai} = \sigma_{ai}^{-}$ when $\xi = -1$, and $\gamma_{ai}\sigma_{ai} = \sigma_{ai}$ when $\xi = 0$. Between

these values, we interpolate. We choose to use a linear interpolation, so that

$$\gamma_{ai}\sigma_{ai} = \begin{cases} (1-\xi)\sigma_{ai} + \xi\sigma_{ai}^{+} & \text{when } \xi > 0 \\ (1-\xi)\sigma_{ai} + \xi\sigma_{ai}^{-} & \text{when } \xi < 0 \end{cases}. \tag{7.14}$$

From this, we can derive an expression for γ_{ai}:

$$\gamma_{ai} = 1 + \xi\frac{\sigma_{ai}^{\pm} - \sigma_{ai}}{\sigma_{ai}}. \tag{7.15}$$

Since we have added several parameters that are not fixed, but are constrained, let us summarize the constraint terms that must be added to the likelihood. Combining the constraint terms for α, β, and ξ, we have

$$\prod_{a,i} \text{Poisson}(m_i|\alpha_i m_i) \cdot \text{Normal}(\xi|1,0) \cdot \text{Normal}(\beta_{ai}|1, S_{ai}) \tag{7.16}$$

7.5.2 Application of the Model

The statistical model is implemented using the RooStats HistFactory tool [71]. We provide to the tool as input the normalized $M_{\text{CT}\perp}$ histograms from the templates outlined in Section 7.3. These specify σ_{ai}. We also provide the shape uncertainties S_{ai} and histogram uncertainty variations $\sigma_{ai}^{\pm}$. The raw counts m_i are determined from the squared sums of weights in the templates.

We construct two different models. The first uses both the same-flavor and opposite-flavor channels; the two channels are combined by multiplying their likelihoods. We use one template for each background category so that $a \in$ (Top, Diboson, Z/γ^*, Non-Prompt). Let us now describe in detail how each template and its systematic uncertainties are derived. Much of this was already described, but we now discuss how everything is implemented in the statistical model.

- **Top:** The template for top-like events is taken from events in data which pass the signal selection, but have at least 1 b-tag. We parameterize the

normalization of the top templates in the two channels so that their ratio is constrained. That is, $N_{of} = rN_{sf}$. The value of r is taken from simulation with a 10% uncertainty due to trigger efficiency uncertainties. Thus, while N_{sf} and N_{of} are both free parameters, we add the constraint $N_{of} \sim \text{Normal}(rN_{sf}, 10\% \times rN_{sf})$ to the likelihood.

- **Diboson and Rare SM:** The diboson template is derived from Monte Carlo simulation, and is the sum of several different components. To create the diboson template, we sum the simulated contributions to the signal region from 6 different processes: WW $\to \ell\nu\ell\nu$, WZ $\to 3\ell\nu$ (the third lepton is missed), ZZ $\to 2\ell 2\nu$, WWZ, WWW (henceforth WWZ and WWW are treated together and called VVV), and H $\to$ WW $\to \ell\nu\ell\nu$. The relative contribution of each background component in this template is fixed to the value predicted by Monte Carlo simulation.

  Like the top template, we parameterize the normalization in the two channels so that their ratio is constrained. That is, $N_{of} = rN_{sf}$. The value of r is taken from simulation with a 10% uncertainty. Thus, while N_{sf} and N_{of} are both free parameters, we add the constraint $N_{of} \sim \text{Normal}(rN_{sf}, 10\% \times rN_{sf})$ to the likelihood.

  There are five histogram systematics applied to this template. These correspond to varying the relative contribution of each background to the template. The $\pm 1\sigma$ shapes are derived by varying the component cross sections up and down by 10% for diboson processes and 50% for triple gauge boson processes and H $\to$ WW. Additionally, there is a shape systematic applied from the comparison of WW Monte Carlo to the $3-1$ lepton control region.

- Z/γ^* **:** This template is also derived from simulation. It is primarily composed by the Drell-Yan process, but also includes contributions from ZZ $\to$

$2\ell 2q$ as this process also has one leptonically decaying Z boson and no real $\not{p}_{\mathrm{T}}$.

The Z/γ^* template uses a shape systematic to describe our imperfect knowledge of the shape of this template. The uncertainty on the bin height is taken from the average fractional discrepancy in each $M_{\mathrm{CT}\perp}$ bin in the control region shown in Figure 7.4, which comes to 24%.

- **Non-prompt:** This template is derived from events which pass the signal selection, but with the isolation of one lepton in a sideband and the two leptons required to have the same, rather than the opposite, charge. The non-prompt template also uses a shape systematic to describe the uncertainty on the shape. We assign a 30% uncertainty to each non-zero bin.
- **Signal:** The signal model for the two-channel analysis is shown in 7.1(a). The template is derived from Monte Carlo simulation. Unlike the other templates, the normalization of the signal template is fixed. Thus, it has several systematic uncertainties on this normalization. We apply Gaussian uncertainties on integrated luminosity (4.4%), trigger efficiency (5%), lepton selection and identification (2%), and b-jet veto efficiency (5%). We also assign a histogram systematic that comes from varying the jet-energy scale up and down by its (jet p_{T}- and η-dependent) uncertainty.

The same-flavor-only analysis uses a slightly different set of templates

- **Flavor-Symmetric:** This template is taken from opposite-flavor events in data which pass all of the signal selections. The template has no systematic uncertainties.
- **Diboson and rare SM:** This template is derived from Monte Carlo simulation. It includes non-flavor symmetric processes: $\mathrm{ZZ} \to 2\ell 2\nu$ and a

WZ $\rightarrow 3\ell\nu$ where the two observed leptons come from the Z boson. The relative contribution of each is fixed to the value in Monte Carlo simulation.

Systematic uncertainties are similar to those in the two-channel analysis, except that in this case, only two processes contribute to this template. We add two histogram systematics which are derived by varying the cross-sections of each component up and down by 10%.

- **Signal:** We interpret the same-flavor-only analysis in terms of the signal process shown in 7.1(b). The systematic uncertainties are the same as in the signal model for the two-channel analysis.

7.5.3 Anderson-Darling Test

Without assuming any particular signal model, we can test how well the data is described by the background-only hypothesis using the Anderson-Darling test [66]. The Anderson-Darling test is used to test the hypothesis that a sample was drawn from a probability distribution. In this purpose, it is similar to the more commonly-used Kolmogorov-Smirnov test. However, the Anderson-Darling test statistic is more sensitive to tail regions of the distribution, making it more useful for searches for new physics, as new physics signals are often expected to appear in the tails of distributions.

The Anderson-Darling statistic is defined in terms of the cumulative distribution function $F(x)$ of the distribution being tested and the empirical distribution function $F_n(x)$ taken from the data sample with n events. It is defined as

$$A^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)[1 - F(x)]} dF(x). \tag{7.17}$$

In this analysis, the probablility distribution being tested is the best-fit sum of background templates. The probability distribution and data are both binned, so

we modify the statistic [72] to

$$A^2 = n \sum_{\text{bins}} \frac{(F_n(x_{\text{hi}}) - F(x_{\text{hi}}))^2}{F(x_{\text{hi}})[1 - F(x_{\text{hi}})]}[(F(x_{\text{hi}}) - F(x_{\text{lo}})], \tag{7.18}$$

where x_{hi} and x_{lo} are the upper and lower edges of the bin.

We also note that Anderson-Darling statistic is additive [73]: two different distributions can be combined by adding their test statistics. We take advantage of this fact to combined the opposite-flavor and same-flavor channels.

There are some cases in which the distribution of the Anderson-Darling statistic is known analytically. The case we deal with here is not one of them. Here, we estimate the distribution of the statistic under the null (background-only) hypothesis using toy Monte Carlo simulations. For each toy experiment, we draw samples from the distribution created by the best fit of the background-only hypothesis. Then we calculate the statistic for each toy by first fitting it to the sum of templates, then calculating A^2 using Equation 7.18. The distribution of A^2 across many toy experiments approximates the distribution of A^2 under the background-only hypothesis. We compare the value of A^2 calculated in data to this distribution to obtain a p-value.

7.5.4 Limit-Setting

In addition to the generic goodness-of-fit test, we set limits on the signal models. Following the procedure outlined in Ref. [74], limit-setting is performed by inverting a hypothesis test which tests the background-only hypothesis against a signal-plus-background hypothesis with the signal cross-section set to some value. We use the CL_S criterion [67], which excludes the signal-plus-background hypothesis at the 95% confidence level (CL) if $\text{CL}_\text{S} < 0.05$. Here,

$$\text{CL}_\text{S} = \frac{p_{s+b}}{1 - p_b}, \tag{7.19}$$

where p_{s+b} and p_b are the p-values of the signal plus background and background-only hypotheses, respectively. Thus, the $\mathrm{CL_S}$ criterion uses the results from two different hypothesis tests: one for the signal-plus-background hypothesis, and one for the background-only hypothesis. This is intended to ensure that the signal-plus-background model is not excluded when the data is consistent with neither the signal-plus-background nor background-only hypothesis. Note that this leads to slightly conservative limits.

The p-values for the signal-plus-background and background-only hypothesis tests are calculated using the profile-likelihood ratio test statistic, as described in Ref. [75]. Denoting the signal cross-section as μ and including all other "nuisance" parameters in the vector θ, we define the test statistic as a function of the observed data $\mathbf{x}$

$$\tilde{q}(\mathbf{x}, \mu) = -2 \ln \frac{\mathcal{L}(\mathbf{x}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\mathbf{x}|\hat{\mu}, \hat{\theta})}. \tag{7.20}$$

Here, $\hat{\theta}_\mu$ is the maximum-likelihood estimate (MLE) for θ when the cross-section is fixed to μ, and $\hat{\mu}$ and $\hat{\theta}$ are the global maximum-likelihood estimates for μ and θ, subject to the constraint $0 < \hat{\mu} < \mu$.

The distributions for the test statistic under both the signal-plus-background and background-only hypotheses are calculated using toy Monte Carlo simulations in which we draw events from the model under the given hypothesis, then calculate the test statistic for these events. To define the distributions that the toy Monte Carlo draws from, for the signal-plus-background hypothesis, we fix μ to the value being tested and θ to its MLE for that value of μ. For the background-only hypothesis, we fix $\mu = 0$, and again θ is fixed to its MLE for $\mu = 0$. After generating the test statistic distributions, it is straightforward to obtain the p-values for the observed data under both hypotheses, and thus the $\mathrm{CL_S}$ value. To set upper limits on a cross-section, we scan over cross-sections to find the smallest

cross-section for which $CL_S \leq 0.05$.

7.5.5 Counting Analysis

As a cross-check to the shape-based analysis, we perform a counting analysis in which we define the high-$M_{CT\perp}$ region as the signal region. We predict the amount of SM background in this region, then check the observed event yield. An excess of events over the predicted number of background events would be interpreted as a sign of new physics.

To make the background prediction, we take the statistical model from the shape analysis and run the background-only fit in the low-$M_{CT\perp}$ region ($10 < M_{CT\perp} < 120$ GeV). This region is expected to have a minimal fraction of signal. We then use use the background normalizations from this signal-free region to extrapolate to the high-$M_{CT\perp}$ signal region ($120 < M_{CT\perp} < 300$ GeV).

Because the top and diboson shapes are very similar in the low-$M_{CT\perp}$ region, the fit has difficulty distinguishing between them, resulting in a biased estimator for the normalizations. We overcome this problem by adding a constraint on the ratio of diboson to top events, which we take from simulation with a 10% systematic uncertainty. This uncertainty is derived from the uncertainty on the CMS $t\bar{t}$, WW, and ZZ cross-section measurements at 8 TeV [76, 77].

7.5.6 Validation

We can validate the statistical model by generating toy experiments from the model for various values of the model parameters. We perform a maximum likelihood fit on the toy experiments with the goal of verifying that the fit returns, on average, the parameters that were used to generate the data.

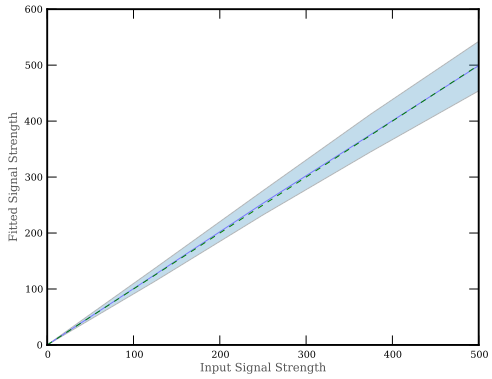
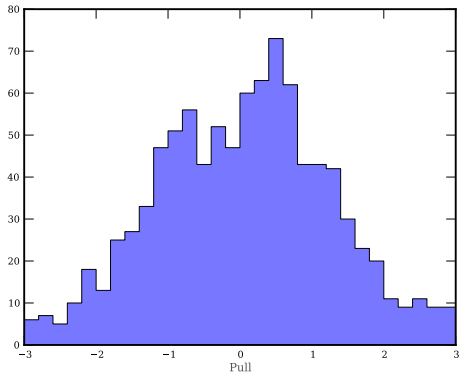


Figure 7.11: Fitted signal strength for toy experiments versus the signal strength used to generate the toys for the mass point $m_{\widetilde{\chi}^{\pm}} = 400$ GeV, $m_{\widetilde{\chi}^{0}_{1}} = 100$ GeV in the chargino pair-production model. The blue line and band indicate the mean and standard deviation over 500 toy experiments per signal strength value. The red dashed line indicates unit slope.

Figure 7.11 shows the median fitted signal strength value over 500 toy experiments per point versus the signal strength value used to generate the toy experiments, where we define signal strength as the cross-section divided by the reference cross-section: $\sigma/\sigma_{\text{ref}}$. The error band gives the central 68% interval of the fitted signal strength over the 500 experiments at each point. The nuisance parameters used to generate the toys are those from the best fit of the observed data with the signal strength floating. These parameters are the same for each toy generated.

Looking at an individual signal strength point, we can look at the distribution of fitted values. We work in terms of a pull variable, defined as (fitted − true)/fit uncertainty. A well-behaved pull distribution will be a unit normal. Figure 7.12 shows the pull distribution over 1000 toys at the mass point $m_{\widetilde{\chi}^{\pm}} = 400$ GeV, $m_{\widetilde{\chi}^{0}_{1}} = 100$ GeV in the chargino pair-production model with a signal strength of 10. The pull mean is 0.0 and the standard deviation is 1.5. We use a large signal strength to keep well away from the lower limit of the signal strength at 0.

We can also check to make sure that the Anderson-Darling test provides dis-



(a)

Figure 7.12: Pull distribution for 1000 toy experiments for the mass point $m_{\widetilde{\chi}^{\pm}} =$ 400 GeV, $m_{\widetilde{\chi}_1^0} = 100$ GeV and signal strength 10 in the chargino pair-production model.

criminating power against signal. Figure 7.13 shows the distribution of the Anderson-Darling test statistic under toy experiments generated from the background-only model and from the signal-plus-background model, where the signal is from the chargino pair-production model with $m_{\widetilde{\chi}^{\pm}} = 400$ GeV, $m_{\widetilde{\chi}_1^0} = 0$ GeV. By comparing the distributions, we can see that there would be a high chance of observing a low p-value if this signal were present.

Finally, we check for biases in the counting analysis by applying the background prediction method on toy Monte Carlo simulations. To do this, we generate toy experiments with two different sets of yields: the ones fitted from data and those from Monte Carlo. For each set of yields, we generate and find the background estimate for 1000 toys. We plot the pull distribution for the background prediction in the signal region. The pull distributions for the fitted and Monte Carlo yields are shown in Figure 7.14 and Figure 7.15, respectively. The means and standard deviations for these distributions are shown in Table 7.2.

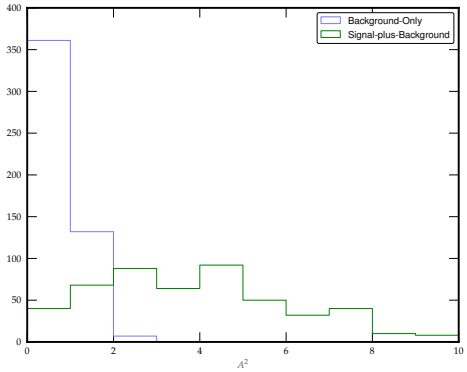
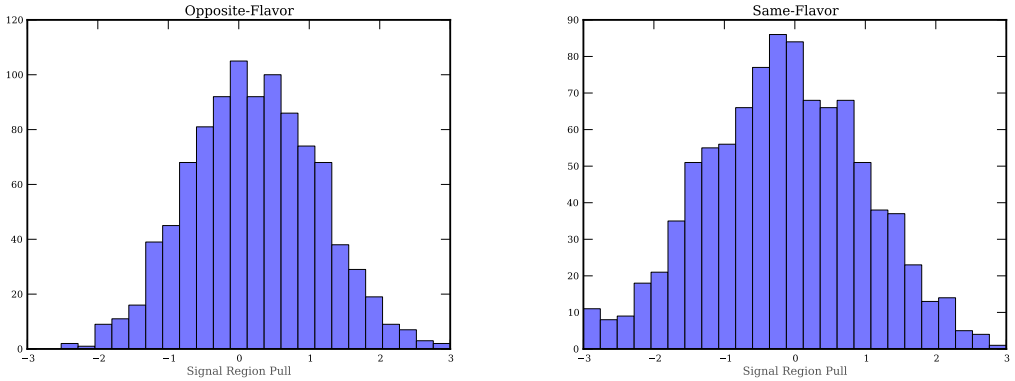


Figure 7.13: Distributions of the Anderson-Darling test statistic for toy experiments generated under the background-only model and the signal-plus-background model for the signal point $m_{\widetilde{\chi}^{\pm}} = 400$ GeV, $m_{\widetilde{\chi}^0_1} = 0$ GeV in the chargino pair-production model.

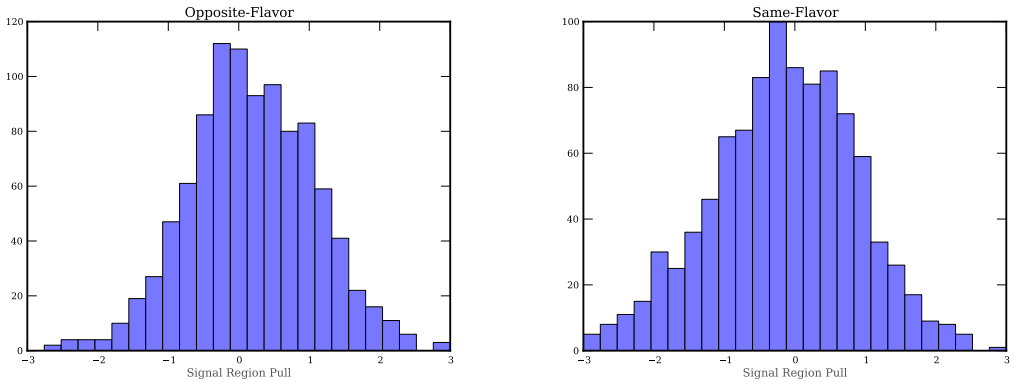


(a) (b)

Figure 7.14: Pull distributions for a) opposite- and b) same-flavor background predictions in the signal region based on toy Monte Carlo samples generated with the yields fitted from data.

Table 7.2: Pull means and standard deviations for the background prediction in the cut-based analysis from toy Monte Carlo experiments.

Input Yield	Channel	Pull Mean	Pull Standard Deviation
Fitted	Same-Flavor	–0.26	1.28
Fitted	Opposite-Flavor	0.21	0.92
Simulated	Same-Flavor	–0.22	1.24
Simulated	Opposite-Flavor	0.19	0.90



(a) (b)

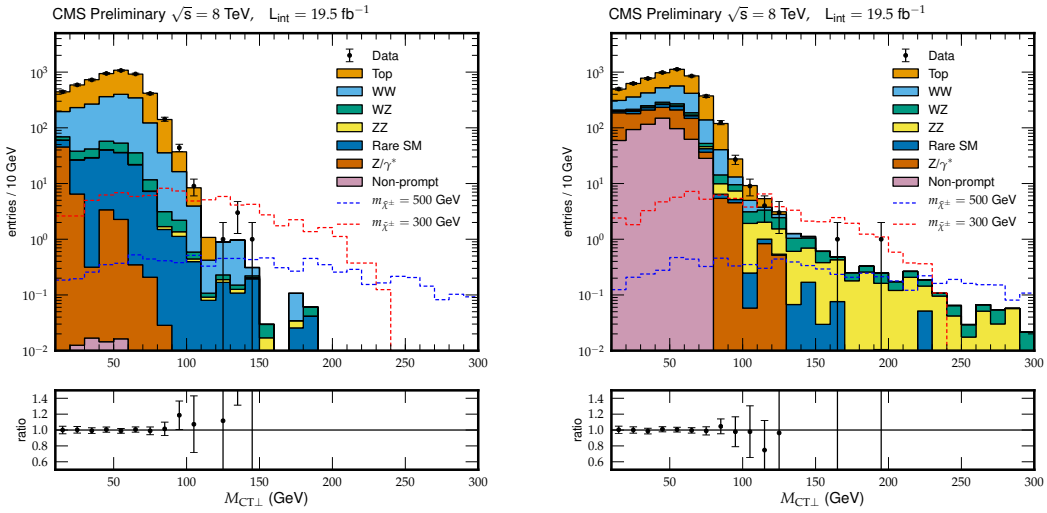
Figure 7.15: Pull distributions for a) opposite- and b) same-flavor background predictions in the signal region based on toy Monte Carlo samples generated with the yields from Monte Carlo simulation.

7.6 Results

7.6.1 Shape Analysis

Figure 7.16 shows the $M_{\mathrm{CT}\perp}$ distributions in data along with the best fits to the sum of background templates. The fitted number of events from each background are displayed in Table 7.3. The data is consistent with the background-only hypothesis; the Anderson-Darling test gives a p-value of 0.55.

Figure 7.17 and Table 7.4 show the results of the fit the the same-flavor channel using opposite-flavor events as the template for flavor-symmetric backgrounds. We see that this fit is consistent with the fit using separate templates. In particular, the number of events from flavor-symmetric backgrounds is about the same. The background-only hypothesis describes the data well, with a p-value of 0.29.



(a) (b)

Figure 7.16: Top: The $M_{\mathrm{CT}\perp}$ distribution in data (black points) and the background prediction (filled histogram) in the opposite-flavor (a) and same-flavor (b) channel. The background shown is a fit to templates derived from control samples or simulation.

Table 7.3: Background yields from the best fit of the background-only hypothesis when fitting both channels using separate templates for top and diboson backgrounds. Also shown for comparison are event yields from Monte Carlo simulation.

	Opposite-Flavor		Same-Flavor	
	Fit	Simulation	Fit	Simulation
Top	3325 ± 561	3071	2581 ± 439	2271
Diboson and Rare SM	1926 ± 178	1466	1599 ± 162	1249
Z/γ^*	58^{+146}_{-58}	106	607 ± 354	920
Non-Prompt	$0.1^{+301}_{-0.1}$	455	603 ± 463	130

Table 7.4: Background yields from the best fit of the background-only hypothesis in the same-flavor channel using opposite-flavor events as a template for flavor symmetric backgrounds. Note that the WZ component here refers only to decays where both identified leptons come from the Z boson. Decays where one lepton comes from the W boson are included in the flavor-symmetric category.

	Fit	Simulation
Flavor-Symmetric	3982 ± 390	3450
Non-FS Diboson	81 ± 42	70
Z/γ^*	381^{+517}_{-381}	920
Non-Prompt	942 ± 627	130

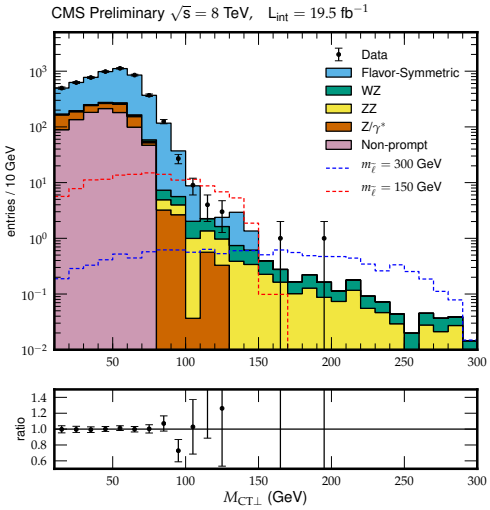


Figure 7.17: The $M_{\mathrm{CT}\perp}$ distribution in data (black points) and the background-only fit (filled histogram) in the same-flavor channel. The background shown is a fit to templates derived from control samples or simulation. For this fit, opposite-flavor events are used to derive a template for the flavor-symmetric backgrounds. Note that the WZ component here refers only to decays where both identified leptons come from the Z boson. Decays where one lepton comes from the W boson are included in the flavor-symmetric category.

7.6.2 Counting Analysis

Table 7.5 shows the results of the counting analysis, in which we fit only in the low $M_{\mathrm{CT}\perp}$ region, then extrapolate to the high $M_{\mathrm{CT}\perp}$ region. We see slightly fewer events than expected in the signal region for the same-flavor channel and more events than expected for the opposite-flavor channel. This excess in the opposite-flavor channel has a local p-value of 0.028, corresponding to a 1.9σ excess. Examining the location of these 5 events in Figure 7.16, we see that they all lie slightly above the $M_{\mathrm{CT}\perp}$ cut, in a region where we would expect some background events. Thus, we attribute this excess to a statistical fluctuation.

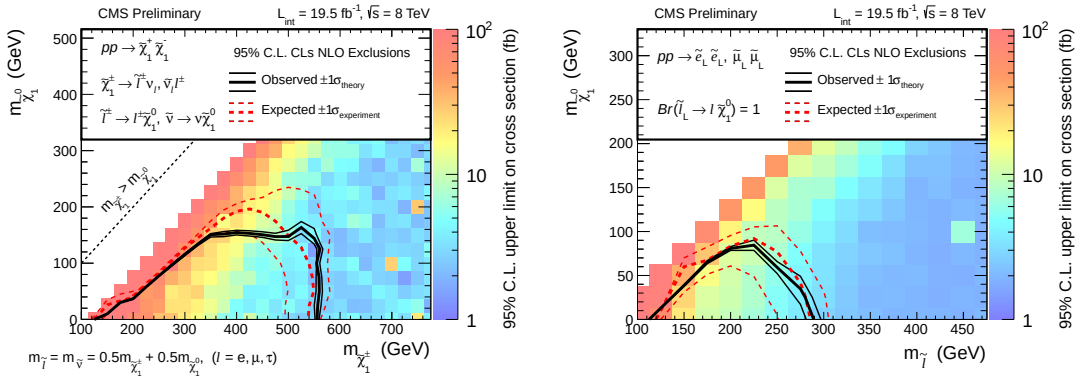
Table 7.5: Low $M_{\mathrm{CT}\perp}$ fit yields and extrapolation to the high $M_{\mathrm{CT}\perp}$ region. Where the predicted value is zero, we give approximate 1σ upper limits.

	Opposite-Flavor		Same-Flavor	
	Low $M_{\mathrm{CT}\perp}$	High $M_{\mathrm{CT}\perp}$	Low $M_{\mathrm{CT}\perp}$	High $M_{\mathrm{CT}\perp}$
Top	3561 ± 152	< 0.4	2704 ± 184	0.3 ± 0.1
Diboson	1678 ± 120	1.6 ± 0.7	1528 ± 140	7.8 ± 1.3
Z/γ^*	67^{+136}_{-67}	< 0.01	601 ± 345	0.5 ± 0.3
Non-Prompt	< 271	< 0.01	555 ± 417	< 0.5
Total	5307 ± 73	1.6 ± 0.8	5389 ± 73	8.7 ± 1.5
Data	5309	5	5388	5

7.6.3 Limits

Figure 7.18 shows the 95% CL upper limits on the cross-section for the two signal models as a function of the particle masses. The exclusion curves show regions where the upper limit is below the reference cross-section; we consider regions below these curves to be excluded, at least for these simple models.

The limits are set using the procedure described in Subsection 7.5.4. For the chargino pair-production model, we use both the opposite- and same-flavor channel, and thus the statistical model which uses separate templates for top and di-



(a) (b)

Figure 7.18: 95% CL upper limits on $\sigma\times$BR for a) the chargino pair-production model and b) the slepton pair-production model.

boson events. For the slepton pair-production model, we use only the same-flavor channel. In this case, we use model that utilizes flavor-symmetric template.

For both models, cross-section upper limits range from a few to hundreds of fb, with more stringent limits for larger mass splittings. This makes sense, as larger mass splittings give larger $M_{\mathrm{CT}\perp}$ values. Also note that there is not much sensitivity near the diagonal, where the charginos or sleptons are nearly degenerate with the LSP. These cases do not yield high $M_{\mathrm{CT}\perp}$ values, and in some cases may not even produce events beyond the SM WW endpoint, so it is no surprise that the analysis cannot set strong limits in these regions.

In the chargino pair-production model, we exclude models with large chargino-LSP mass splittings up to chargino masses of around 550 GeV. This is in line with the expected exclusions under the no-signal hypothesis, which are shown by the red lines on the figure. For the slepton pair-production model, the exclusion goes up to only slepton masses of around 275 GeV, again in agreement with the expceted limits. The difference in reach between the two models is largely due to the much smaller reference cross-section for slepton-pair production compared to chargino pair-production.

7.7 Conclusion

We performed an analysis that was based on the shape of the $M_{\mathrm{CT}\perp}$ distribution in events with two leptons and $\not{p}_{\mathrm{T}}$. By considering only the shapes of the distributions, we were able to focus on deriving the shapes of the $M_{\mathrm{CT}\perp}$ distributions for various background processes, while leaving the normalizations of these processes as free parameters. This search is sensitive to weakly-produced charginos and sleptons.

We presented several different results. First, we presented a generic goodness-of-fit p-value for the combination of opposite-flavor and same-flavor lepton channels and for the same-flavor channel by itself. We found that the data fit well to the background-only hypothesis. We performed a count-based analysis, which predicted the number of background events in the high-$M_{\mathrm{CT}\perp}$ region by extrapolating from the low-$M_{\mathrm{CT}\perp}$ region. We again found good agreement with SM background predictions. Finally, we used a binned profile-likelihood ratio-based approach to set limits on two different signal models: one describing chargino pair-production with decays via sleptons and sneutrinos, and the other describing slepton pair production.

We have demonstrated that kinematic endpoint variables are extremely useful in searches for new physics. Their utility arises because many SM backgrounds have definite upper endpoints. Thus by exploring the region above the endpoint, we can gain sensitivity to new physics.

CHAPTER 8

CONCLUSION

While there has so far been no sign of physics beyond the Standard Model at the LHC, there are excellent reasons to believe that such physics must exist. The LHC is currently shut down for an upgrade to around 13 TeV center of mass energy. This increase in energy will once again open up a new region to explore in the search for new physics.

This work has focussed on exploring and developing a set of tools that will be useful in both searches for new physics and in characterization of new physics when and if it is found. We showed that the experimental issues that arise in applying kinematic endpoint variables to mass measurements can be overcome. We also demonstrated an alternate use for these variables by using them in a search for new physics in events with two leptons and $\not{p}_{\mathrm{T}}$.

In addition to demonstrating the utility of kinematic endpoint variables, we have achieved two nice results that are contributions to the field of particle physics in their own right. We presented a measurement of the top-quark mass that is very different from traditional methods and has similar precision. The fact that the result is very similar to previous results gives confidence in these measurements.

We also probed a region that has not been thoroughly explored at the LHC: weakly produced new physics with little or no hadronic activity. We set what are currently the world's best limits on selectron, smuon, and chargino pair-production. These results could also be extended to other theories in which particles are pair produced, then decay into leptons and $\not{p}_{\mathrm{T}}$.

There is plenty of room for future work involving kinematic endpoint variables. Here, we have used only two of the many variables and techniques that were described in Chapter 5. In addition, kinematic endpoint variables remain an area

of active research in the phenomenology community.

In addition to this broader exploration of new techniques, both analyses presented here leave some room for future work. Of course, both analyses will benefit from more data collected by CMS. In the top-quark mass measurement, one of the dominant systematic uncertainties comes from the choice of fit range. Work to reduce this uncertainty would greatly increase the precision of the measurement.

The major weakness of the analysis presented in Chapter 7 is its lack of sensitivity at small mass splittings. This is in part fundamental to the analysis, but the sensitivity could potentially be improved by fixing some or all of the background normalizations, rather than fitting for them. Improvements might also give sensitivity to chargino decays via W bosons, which is currently not available due to the small branching ratio of $\mathrm{WW} \to \ell\nu\ell\nu$.

In conclusion, we have shown that kinematic endpoint variables *are* useful when applied to real data from a real detector. They are not just an invention that works well in theory. They have a real place as an experimental technique, and can contribute to understanding the solutions to the problems that are currently facing particle physics.

BIBLIOGRAPHY

[1] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". *Phys.Lett.* B716 (2012), 30–61. arXiv: 1207.7235 [hep-ex].

[2] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Phys.Lett.* B716 (2012), 1–29. arXiv: 1207.7214 [hep-ex].

[3] J. Wess and B. Zumino. "Supergauge Transformations in Four-Dimensions". *Nucl.Phys.* B70 (1974), 39–50.

[4] S. P. Martin. "A Supersymmetry primer" (1997). arXiv: hep-ph/9709356 [hep-ph].

[5] S. Dimopoulos and D. W. Sutter. "The Supersymmetric flavor problem". *Nucl.Phys.* B452 (1995), 496–512. arXiv: hep-ph/9504415 [hep-ph].

[6] M. Papucci, J. T. Ruderman, and A. Weiler. "Natural SUSY Endures". *JHEP* 1209 (2012), 035. arXiv: 1110.6926 [hep-ph].

[7] O. S. Bruning et al. "LHC Design Report. 1. The LHC Main Ring". CERN-2004-003-V-1, CERN-2004-003 (2004).

[8] CMS Collaboration. "CMS physics: Technical design report". CERN-LHCC-2006-001, CMS-TDR-008-1 (2006).

[9] CMS Collaboration. "The CMS experiment at the CERN LHC". *JINST* 3 (2008), S08004.

[10] CMS Collaboration. "CMS, tracker technical design report". CERN-LHCC-98-06, CMS-TDR-5 (1998).

[11] CMS Collaboration. "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays". *JINST* 5 (2010), T03021. arXiv: 0910.5530 [physics.ins-det].

[12] CMS Collaboration. "CMS: The electromagnetic calorimeter. Technical design report". CERN-LHCC-97-33 (1997).

[13] CMS Trigger and Data Acquisition Group. "The CMS high level trigger". *Eur.Phys.J.* C46 (2006), 605–667. arXiv: hep-ex/0512077 [hep-ex].

[14] R. E. Kalman et al. "A new approach to linear filtering and prediction problems". *Journal of Basic Engineering* 82.1 (1960), 35–45.

[15] R. Fruhwirth. "Track fitting with long tailed noise: A Bayesian approach". *Comput.Phys.Commun.* 85 (1995), 189–199.

[16] R. Fruhwirth. "Track fitting with nonGaussian noise". *Comput.Phys.Commun.* 100 (1997), 1–16.

[17] CMS Collaboration. "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET". CMS-PAS-PFT-09-001 (2009).

[18] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_t jet clustering algorithm". *JHEP* 04 (2008), 063. arXiv: 0802.1189 [hep-ex].

[19] CMS Collaboration. "Missing transverse energy performance of the CMS detector". *JINST* 6 (2011), P09001. arXiv: 1106.5048 [physics.ins-det].

[20] A. J. Barr and C. G. Lester. "A Review of the Mass Measurement Techniques proposed for the Large Hadron Collider". *J.Phys.* G37 (2010), 123001. arXiv: 1004.2732 [hep-ph].

[21] UA1 Collaboration. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV". *Phys.Lett.* B122 (1983), 103–116.

[22] UA2 Collaboration. "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider". *Phys.Lett.* B122 (1983), 476–485.

[23] C. Lester and D. Summers. "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders". *Phys.Lett.* B463 (1999), 99–103. arXiv: hep-ph/9906349 [hep-ph].

[24] A. Barr, C. Lester, and P. Stephens. "m(T2): The Truth behind the glamour". *J.Phys.* G29 (2003), 2343–2363. arXiv: hep-ph/0304226 [hep-ph].

[25] H.-C. Cheng and Z. Han. "Minimal Kinematic Constraints and m(T2)". *JHEP* 0812 (2008), 063. arXiv: 0810.5178 [hep-ph].

[26] C. G. Lester. "Model independent sparticle mass measurements at ATLAS". PhD thesis. Geneva: Cambridge Univ., 2001.

[27] W. S. Cho et al. "Measuring superparticle masses at hadron collider using the transverse mass kink". *JHEP* 0802 (2008), 035. arXiv: 0711.4526 [hep-ph].

[28] M. Burns et al. "Using Subsystem MT2 for Complete Mass Determinations in Decay Chains with Missing Energy at Hadron Colliders". *JHEP* 0903 (2009), 143. arXiv: 0810.5576 [hep-ph].

[29] A. J. Barr, B. Gripaios, and C. G. Lester. "Weighing Wimps with Kinks at Colliders: Invisible Particle Mass Measurements from Endpoints". *JHEP* 0802 (2008), 014. arXiv: 0711.4008 [hep-ph].

[30] K. T. Matchev et al. "Precision sparticle spectroscopy in the inclusive same-sign dilepton channel at LHC". *Phys.Rev.* D82 (2010), 077701. arXiv: 0909.4300 [hep-ph].

[31] W. S. Cho et al. "Gluino Stransverse Mass". *Phys.Rev.Lett.* 100 (2008), 171801. arXiv: 0709.0288 [hep-ph].

[32] D. R. Tovey. "On measuring the masses of pair-produced semi-invisibly decaying particles at hadron colliders". *JHEP* 0804 (2008), 034. arXiv: 0802.2879 [hep-ph].

[33] M. Serna. "A Short comparison between m(T2) and m(CT)". *JHEP* 0806 (2008), 004. arXiv: 0804.3344 [hep-ph].

[34] G. Polesello and D. R. Tovey. "Supersymmetric particle mass measurement with the boost-corrected contransverse mass". *JHEP* 1003 (2010), 030. arXiv: 0910.0174 [hep-ph].

[35] K. T. Matchev and M. Park. "A General method for determining the masses of semi-invisibly decaying particles at hadron colliders". *Phys.Rev.Lett.* 107 (2011), 061801. arXiv: 0910.1584 [hep-ph].

[36] P. Konar et al. "Superpartner Mass Measurement Technique using 1D Orthogonal Decompositions of the Cambridge Transverse Mass Variable M_{T2}". *Phys.Rev.Lett.* 105 (2010), 051802. arXiv: 0910.3679 [hep-ph].

[37] CDF Collaboration. "Top Quark Mass Measurement using mT2 in the Dilepton Channel at CDF". *Phys.Rev.* D81 (2010), 031102. arXiv: 0911.2956 [hep-ex].

[38] CDF Collaboration. "Top quark mass measurement using the template method at CDF". *Phys.Rev.* D83 (2011), 111101. arXiv: 1105.0192 [hep-ex].

[39] LHC/LC Study Group. "Physics interplay of the LHC and the ILC". *Phys.Rept.* 426 (2006), 47–358. arXiv: hep-ph/0410364 [hep-ph].

[40] CMS Collaboration. "Search for supersymmetry in hadronic final states using MT2 in pp collisions at $\sqrt{s} = 7$ TeV". *JHEP* 1210 (2012), 018. arXiv: 1207.1798 [hep-ex].

[41] ATLAS Collaboration. "Search for scalar bottom pair production with the ATLAS detector in pp Collisions at $\sqrt{s} = 7$ TeV". *Phys.Rev.Lett.* 108 (2012), 181802. arXiv: 1112.3832 [hep-ex].

[42] ATLAS Collaboration. "Search for direct slepton and gaugino production in final states with two leptons and missing transverse momentum with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV". *Phys.Lett.* B718 (2013), 879–901. arXiv: 1208.2884 [hep-ex].

[43] W. S. Cho et al. "Measuring the top quark mass with m(T2) at the LHC". *Phys.Rev.* D78 (2008), 034019. arXiv: 0804.2185 [hep-ph].

[44] CDF Collaboration. "First simultaneous measurement of the top quark mass in the lepton + jets and dilepton channels at CDF". *Phys.Rev.* D79 (2009), 092005. arXiv: 0809.4808 [hep-ex].

[45] CMS Collaboration. "Measurement of the top-quark mass in $t\bar{t}$ events with dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV". *Eur.Phys.J.* C72 (2012), 2202. arXiv: 1209.2393 [hep-ex].

[46] D0 Collaboration. "Precise measurement of the top quark mass in the dilepton channel at D0". *Phys.Rev.Lett.* 107 (2011), 082004. arXiv: 1105.0320 [hep-ex].

[47] CMS Collaboration. "Identification of b-quark jets with the CMS experiment". *JINST* 8 (2013), P04013. arXiv: 1211.4462 [hep-ex].

[48] S. Frixione and B. R. Webber. "Matching NLO QCD computations and parton shower simulations". *JHEP* 0206 (2002), 029. arXiv: hep-ph/0204244 [hep-ph].

[49] T. Sjostrand, S. Mrenna, and P. Z. Skands. "PYTHIA 6.4 Physics and Manual". *JHEP* 0605 (2006), 026. arXiv: hep-ph/0603175 [hep-ph].

[50] J. Alwall et al. "MadGraph 5 : Going Beyond". *JHEP* 1106 (2011), 128. arXiv: 1106.0522 [hep-ph].

[51] GEANT4. "GEANT4: A Simulation toolkit". *Nucl.Instrum.Meth.* A506 (2003), 250–303.

[52] B. W. Silverman. *Density estimation for statistics and data analysis*. Vol. 26. Chapman & Hall/CRC, 1986.

[53] D. W. Scott. *Multivariate density estimation: theory, practice, and visualization*. Vol. 383. Wiley, 2009.

[54] W. T. Eadie and F. James. *Statistical methods in experimental physics*. World Scientific Publishing Company, 2006.

[55] B. Efron and R. Tibshirani. *An introduction to the bootstrap*. Vol. 57. Chapman & Hall/CRC, 1993.

[56] A. J. Canty and A. C. Davison. "Resampling-based Variance Estimation for Labour Force Surveys". *Journal of the Royal Statistical Society: Series D (The Statistician)* 48.3 (1999), 379–391. ISSN: 1467-9884. URL: http://dx.doi.org/10.1111/1467-9884.00196.

[57] F. James et al. *Minuit2*. June 2013. URL: http://lcgapp.cern.ch/project/cls/work-packages/mathlibs/minuit/index.html.

[58] R. Barlow. "Application of the bootstrap resampling technique to particle physics experiments". Manchester University Preprint, MAN/HEP/99/4. 1999. URL: http://www.hep.man.ac.uk/preprints/1999.html.

[59] Particle Data Group. "Review of Particle Physics (RPP)". *Phys.Rev.* D86 (2012), 010001.

[60] CMS Collaboration. "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS". *JINST* 6 (2011), P11002. arXiv: 1107.4277 [physics.ins-det].

[61] P. Z. Skands. "Tuning Monte Carlo Generators: The Perugia Tunes". *Phys.Rev.* D82 (2010), 074018. arXiv: 1005.3457 [hep-ph].

[62] E. Lytken. "Prospects for slepton searches with ATLAS". PhD thesis. Copenhagen: Niels Bohr Institutet, Oct. 2003.

[63] M. M. Nojiri et al. "Inclusive transverse mass analysis for squark and gluino mass determination". *JHEP* 0806 (2008), 035. arXiv: 0802.2412 [hep-ph].

[64] A. J. Barr and C. Gwenlan. "The Race for supersymmetry: Using m(T2) for discovery". *Phys.Rev.* D80 (2009), 074007. arXiv: 0907.2713 [hep-ph].

[65] LHC New Physics Working Group. "Simplified Models for LHC New Physics Searches". *J.Phys.* G39 (2012), 105005. arXiv: 1105.2838 [hep-ph].

[66] T. Anderson and D. Darling. "Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes". *The Annals of Mathematical Statistics* 23.2 (1952), 193–212.

[67] T. Junk. "Confidence level computation for combining searches with small statistics". *Nucl.Instrum.Meth.* A434 (1999), 435–443. arXiv: hep-ex/9902006 [hep-ex].

[68] W. Beenakker, R. Hopker, and M. Spira. "PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD" (1996). arXiv: hep-ph/9611232 [hep-ph].

[69] CMS Collaboration. "The fast simulation of the CMS detector at LHC". *J.Phys.Conf.Ser.* 331 (2011), 032049.

[70] R. J. Barlow and C. Beeston. "Fitting using finite Monte Carlo samples". *Comput.Phys.Commun.* 77 (1993), 219–228.

[71] K. Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. CERN-OPEN-2012-016. New York: New York U., Jan. 2012.

[72] L. Demortier. *Assessing the Significance of a Deviation in the Tail of a Distribution*. CDF Note 3419. 1995.

[73] F. Scholz and M. Stephens. "K-sample Anderson–Darling tests". *Journal of the American Statistical Association* 82.399 (1987), 918–924.

[74] ATLAS and CMS Collaborations. "Procedure for the LHC Higgs boson search combination in summer 2011". ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005 (2011).

[75] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". *Eur.Phys.J.* C71 (2011), 1554. arXiv: 1007.1727 [physics.data-an].

[76] *Measurement of the differential ttbar cross section in the dilepton channel at 8 TeV*. Tech. rep. CMS-PAS-TOP-12-028. Geneva: CERN, 2013.

[77] CMS Collaboration. "Measurement of W+W- and ZZ production cross sections in pp collisions at sqrt(s) = 8 TeV". *Phys.Lett.* B721 (2013), 190–211. arXiv: 1301.4698 [hep-ex].